

SOEP Survey Papers

Series C – Data Documentation

SOEP – The German Socio-Economic Panel study at DIW Berlin

2017

Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2015)

Martin Kroh, Simon Kühne, Rainer Siegers

Running since 1984, the German Socio-Economic Panel study (SOEP) is a wide-ranging representative longitudinal study of private households, located at the German Institute for Economic Research, DIW Berlin.

The aim of the SOEP Survey Papers Series is to thoroughly document the survey's data collection and data processing. The SOEP Survey Papers is comprised of the following series:

Series A – Survey Instruments (Erhebungsinstrumente)

Series B – Survey Reports (Methodenberichte)

Series C – Data Documentation (Datendokumentationen)

Series D – Variable Descriptions and Coding

Series E – SOEPmonitors

Series F – SOEP Newsletters

Series G – General Issues and Teaching Materials

The SOEP Survey Papers are available at <http://www.diw.de/soepsurveypapers>

Editors:

Dr. Jan Goebel, DIW Berlin

Prof. Dr. Martin Kroh, DIW Berlin and Humboldt Universität Berlin

Prof. Dr. Carsten Schröder, DIW Berlin and Freie Universität Berlin

Prof. Dr. Jürgen Schupp, DIW Berlin and Freie Universität Berlin

Please cite this paper as follows:

Martin Kroh, Simon Kühne, Rainer Siegers. 2017. Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2015). SOEP Survey Papers 408: Series C. Berlin: DIW/SOEP



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

© 2017 by SOEP

ISSN: 2193-5580 (online)

DIW Berlin

German Socio-Economic Panel (SOEP)

Mohrenstr. 58

10117 Berlin

Germany

soeppapers@diw.de



Data Documentation:

Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2015)

Martin Kroh Simon Kühne Rainer Siegers

November 10, 2016

DIW Berlin – Deutsches Institut für Wirtschaftsforschung e.V.
Sozio-oekonomisches Panel (SOEP)
Mohrenstraße 58
10117 Berlin

Contents

1	Introduction	7
2	Developments in Sample Size	8
2.1	Development of the Number of Successful Interviews by Cross-Section	8
2.2	Continuance and Exit: The First Wave Gross Samples and their Participatory Behavior	22
2.3	New Entrants through Birth or Move into SOEP Households and Their Participation Behavior	30
2.4	The Risk of Survey-Related Panel Attrition	38
3	Panel Attrition Due to Unsuccessful Follow-Ups	43
3.1	The Frequency of Successful Follow-Ups	43
3.2	Predicting the Probability of Successful vs. Unsuccessful Follow-Ups in the Year 2015	45
4	Panel Attrition Due to Refusals	49
4.1	The Frequency of Participation	49
4.2	Predicting the Probability of Re-Interviewing versus Refusal in the Year 2015	51
5	Summary Statistics of the Derived Longitudinal and Cross-Sectional Weights	60

List of Tables

3.1	The Frequency of Households to be Re-Contacted and the Percentage of Successful Follow-Ups, Subsamples A to M1 by Year.	44
3.2	Definition of the Regressors of the Logit Model of Unsuccessful Follow-Ups	46
3.3	Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2015	47
4.1	The Frequency of Re-Contacted Households and the Percentage of Participation, Subsamples A to M by Year.	50
4.2	Definition of the Regressors of the Logit Model of Refusal	52
4.3	Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2015	56
5.1	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples A through D (Percentiles of \$HBLEIB up to Wave BF).	61
5.2	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples E through G (Percentiles of \$HBLEIB up to Wave BF).	62
5.3	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples H, J and K (Percentiles of \$HBLEIB up to Wave BF).	63
5.4	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples L1, L2, L3 and M1 (Percentiles of \$HBLEIB up to Wave BF).	63
5.5	Summary Statistics of the Derived Cross-Sectional Weights at the Household Level (Percentiles of \$HHRF up to Wave 32).	64
5.6	Summary Statistics of the Derived Cross-Sectional Weights at the Person Level (Percentiles of \$PHRF up to Wave 32).	65

List of Figures

1	The Number of Successful Interviews with Persons by Subsamples A through M2, Waves 1 to 32.	8
2	Comparison of Successful Interviews with Persons and Households (Subsamples A and B), Waves 1 to 32	9
3	Comparison of Successful Interviews with Persons and Households (Subsamples C), Waves 1 to 26	10
4	Comparison of Successful Interviews with Persons and Households (Subsamples D), Waves 1 to 21	11
5	Comparison of Successful Interviews with Persons and Households (Subsamples E), Waves 1 to 18 ¹	12
6	Comparison of Successful Interviews with Persons and Households (Subsamples F), Waves 1 to 16	13
7	Comparison of Successful Interviews with Persons and Households (Subsamples G), Waves 1 to 14	14
8	Comparison of Successful Interviews with Persons and Households (Subsamples H), Waves 1 to 10	15
9	Comparison of Successful Interviews with Persons and Households (Subsample J), Waves 1 to 5	16
10	Comparison of Successful Interviews with Persons and Households (Subsample K), Waves 1 to 4	17
11	Comparison of Successful Interviews with Persons and Households (Subsample L1), Waves 1 to 6	18
12	Comparison of Successful Interviews with Persons and Households (Subsample L2), Waves 1 to 6	19
13	Comparison of Successful Interviews with Persons and Households (Subsample L3), Waves 1 to 5	20
14	Comparison of Successful Interviews with Persons and Households (Subsample M1), Waves 1 to 3	21
15	All First-Wave Persons (Gross Subsample A). Development up to Wave 32	22
16	All First-Wave Persons (Gross Subsample B). Development up to Wave 32	23
17	All First-Wave Persons (Gross Subsample C). Development up to Wave 26	23
18	All First-Wave Persons (Gross Subsample D). Development up to Wave 21	24
19	All First-Wave Persons (Gross Subsample E). Development up to Wave 18	24
20	All First-Wave Persons (Gross Subsample F). Development up to Wave 16	25
21	All First-Wave Persons (Gross Subsample G). Development up to Wave 14	25
22	All First-Wave Persons (Gross Subsample H). Development up to Wave 10	26
23	All First-Wave Persons (Gross Subsample J). Development up to Wave 5	26
24	All First-Wave Persons (Gross Subsample K). Development up to Wave 4	27
25	All First-Wave Persons (Gross Subsample L1). Development up to Wave 6	27
26	All First-Wave Persons (Gross Subsample L2). Development up to Wave 6	28
27	All First-Wave Persons (Gross Subsample L3). Development up to Wave 5	28
28	All First-Wave Persons (Gross Subsample M1). Development up to Wave 3	29
29	Entrants and their Participation Behavior (Subsample A). Development up to Wave 32	30

30	Entrants and their Participation Behavior (Subsample B). Development up to Wave 32	31
31	Entrants and their Participation Behavior (Subsample C). Development up to Wave 26	31
32	Entrants and their Participation Behavior (Subsample D). Development up to Wave 21	32
33	Entrants and their Participation Behavior (Subsample E). Development up to Wave 18	32
34	Entrants and their Participation Behavior (Subsample F). Development up to Wave 16	33
35	Entrants and their Participation Behavior (Subsample G). Development up to Wave 14	33
36	Entrants and their Participation Behavior (Subsample H). Development up to Wave 10	34
37	Entrants and their Participation Behavior (Subsample J). Development up to Wave 5	34
38	Entrants and their Participation Behavior (Subsample K). Development up to Wave 4	35
39	Entrants and their Participation Behavior (Subsample L1). Development up to Wave 6	35
40	Entrants and their Participation Behavior (Subsample L2). Development up to Wave 6	36
41	Entrants and their Participation Behavior (Subsample L3). Development up to Wave 5	36
42	Entrants and their Participation Behavior (Subsample M1). Development up to Wave 3	37
43	Successful Re-Interviewing of First-Wave Respondents by Subsamples A, B, C. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	38
44	Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	39
45	Successful Re-Interviewing of First-Wave Respondents by Subsamples G, H, J and K. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	39
46	Successful Re-Interviewing of First-Wave Respondents by Subsamples L1, L2, L3 and M1. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	40
47	Successful Re-Interviewing of All First-Wave Respondents by Age Categories. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	40
48	Successful Re-Interviewing of All First-Wave Respondents by Occupation. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	41

49	Successful Re-Interviewing of All First-Wave Respondents by Income Quintiles. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	41
50	Successful Re-Interviewing of All First-Wave Respondents by Education. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	42

1 Introduction

This data documentation¹ is meant to provide SOEP users with a general overview of the longitudinal development of the survey over the past 32 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 K individually, the boost samples of specific family types L1-L3, the 2013 IAB-SOEP Migration Sample M1 and the recently integrated 2015 IAB-SOEP Migration Sample M2. For a general overview on the integration of refreshment samples into the SOEP see Kroh et al. (2015).

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. (2010). 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); Schonlau et al. (2013) and for a general overview, Haisken-DeNew and 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 2015 based on the characteristics of households measured in 2014. 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 2015: $B_{FHBLEIB}$ and $B_{FPBLEIB}$. Based on the inverse of the probability of observing households and persons in 2014, the staying probability in 2015, and additional post-stratification to meet benchmarks of known marginals of the underlying population in 2015, we derive the cross-sectional weights B_{FHHRF} and B_{FPHRF} . 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.

¹We would like to thank Veronika Belcheva and Myriel Ravagli for their help and contributions.

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 among the 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. This percentage 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 first wave 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

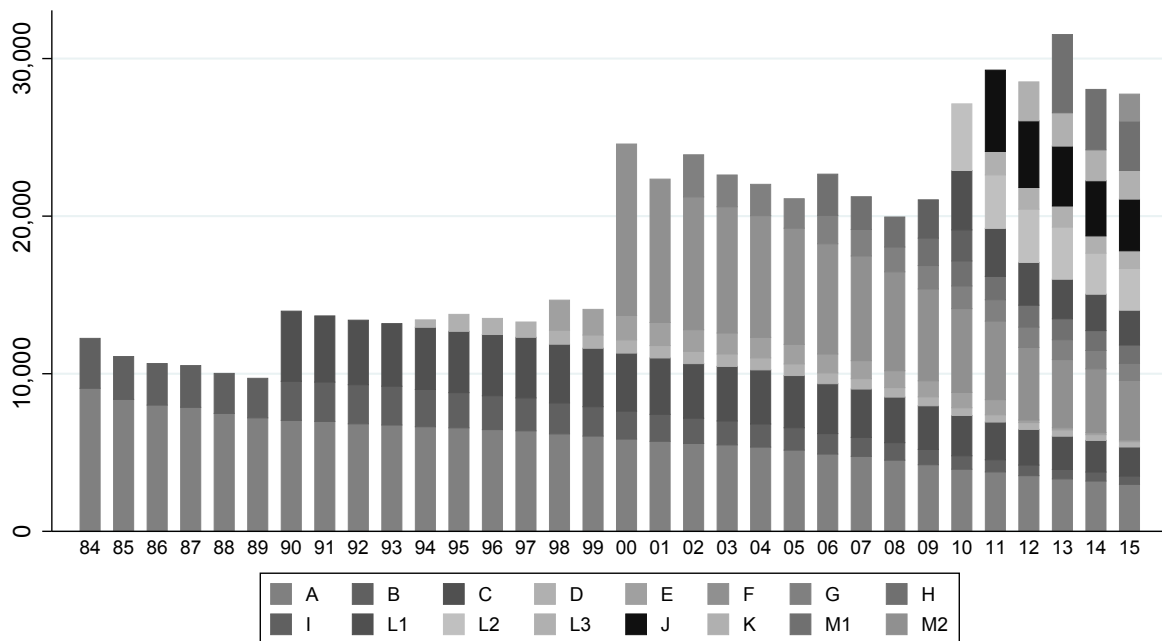


Figure 1: The Number of Successful Interviews with Persons by Subsamples A through M2, Waves 1 to 32.

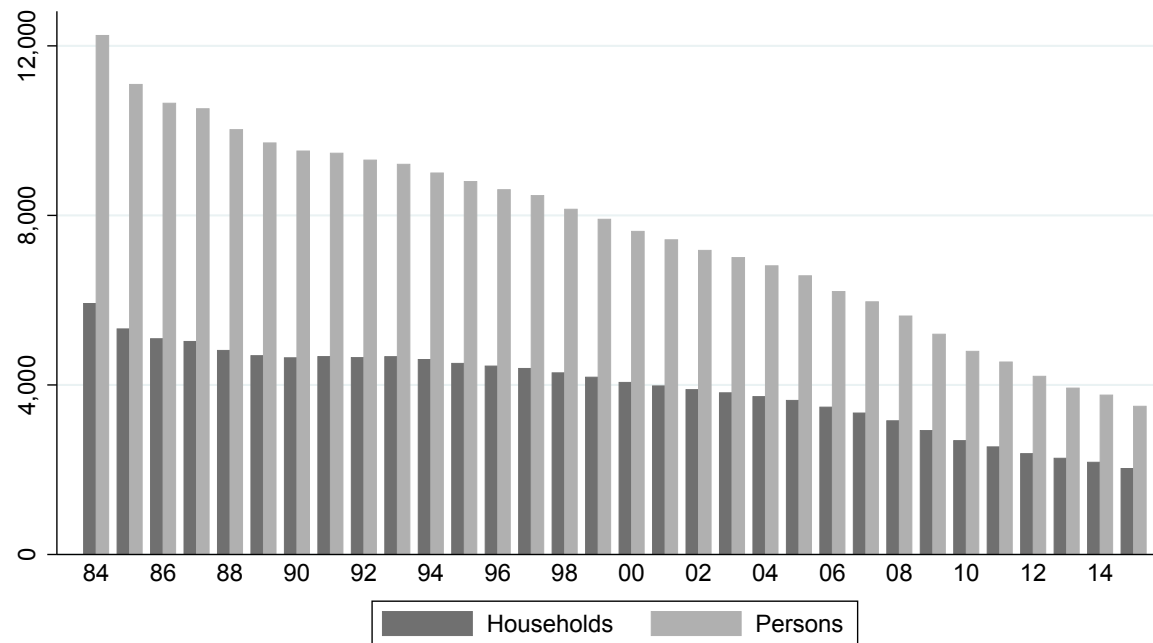


Figure 2: Comparison of Successful Interviews with Persons and Households (Subsamples A and B), Waves 1 to 32

Year	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Persons	7.909	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
Households	4.183	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

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Persons	7.623	7.424	7.175	7.004	6.811	6.575	6.203	5.961	5.626	5.197	4.793	4.541	4.204	3.926	3.761	3.497
Households	4.060	3.977	3.889	3.814	3.724	3.635	3.476	3.337	3.154	2.923	2.686	2.539	2.379	2.270	2.176	2.028

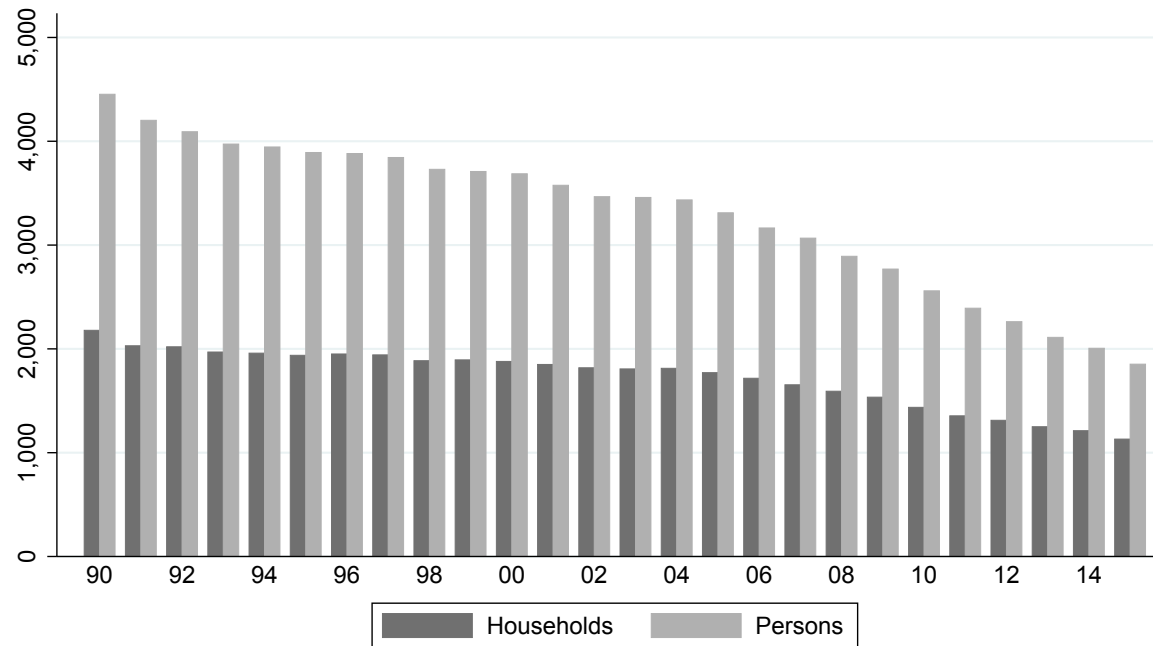


Figure 3: Comparison of Successful Interviews with Persons and Households (Subsamples C), Waves 1 to 26

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
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
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

Year	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Persons	3.459	3.435	3.311	3.165	3.067	2.892	2.769	2.559	2.392	2.262	2.111	2.006	1.853
Households	1.807	1.813	1.771	1.717	1.654	1.592	1.535	1.437	1.355	1.312	1.250	1.212	1.131

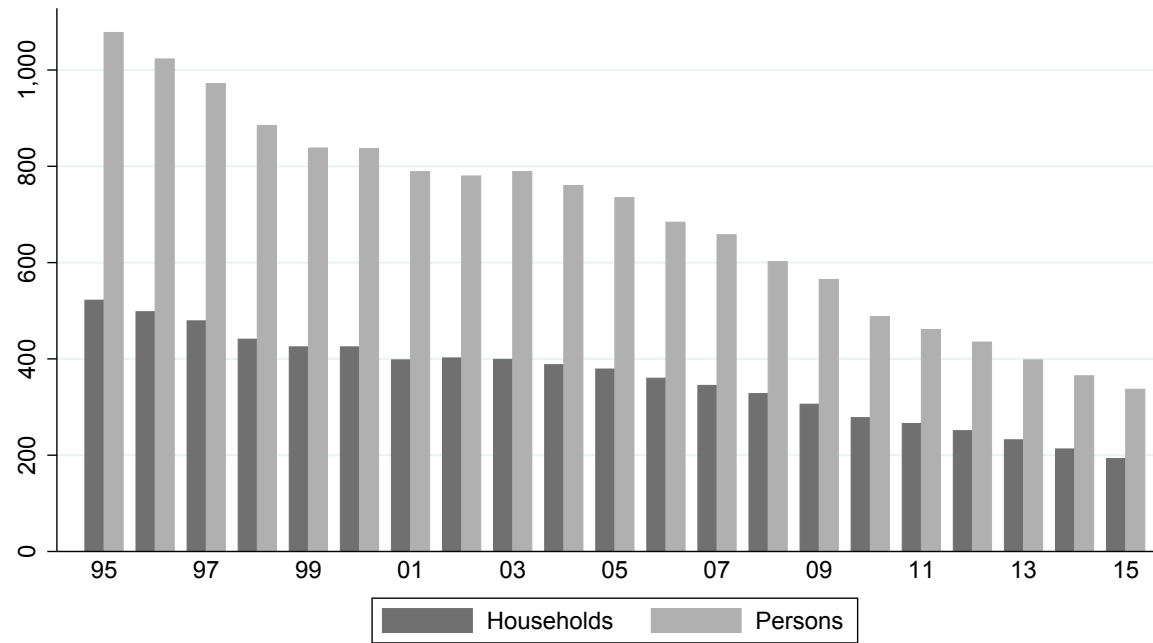


Figure 4: Comparison of Successful Interviews with Persons and Households (Subsamples D), Waves 1 to 21

Year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Persons	1.078	1.023	972	885	838	837	789	780	789	760	735	684	658	602	565	488	461	435	398	365	337
Households	522	498	479	441	425	425	398	402	399	388	379	360	345	328	306	278	266	251	232	213	193

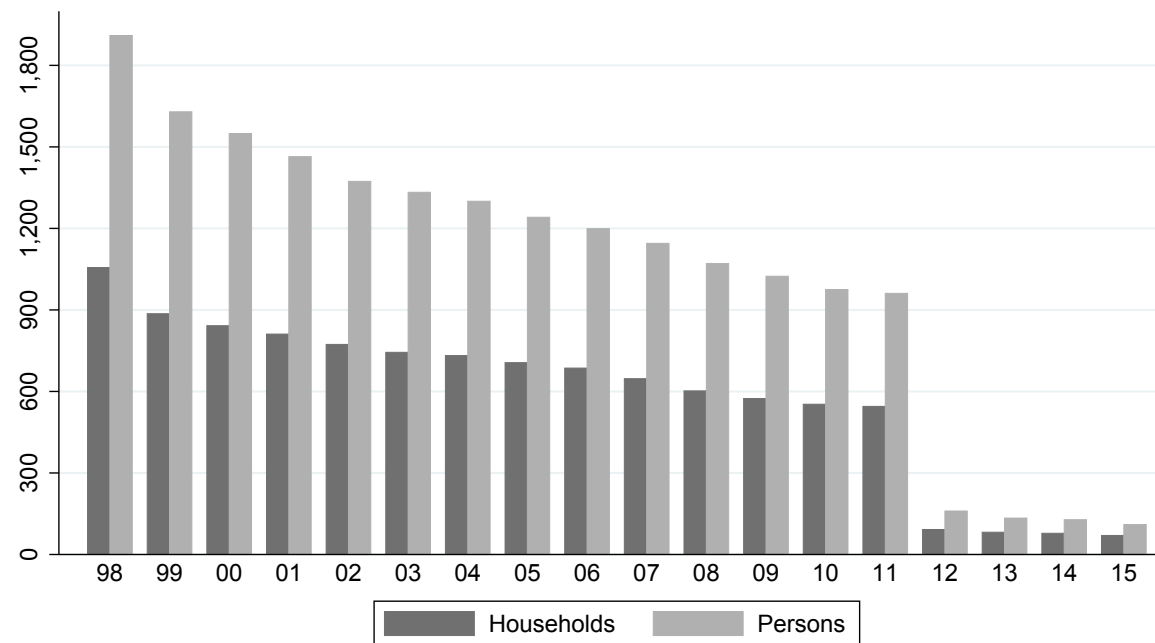


Figure 5: Comparison of Successful Interviews with Persons and Households (Subsamples E), Waves 1 to 18*

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Persons	1.910	1.629	1.549	1.464	1.373	1.333	1.300	1.241	1.199	1.145	1.071	1.024	975	961	160	134	128	110
Households	1.056	886	842	811	773	744	732	706	686	647	602	574	553	545	92	82	78	70

*In 2012, Subsample E has been split into two parts, one being surveyed continuously by SOEP-Core and the larger part being surveyed by SOEP-IS since 2012 onwards.

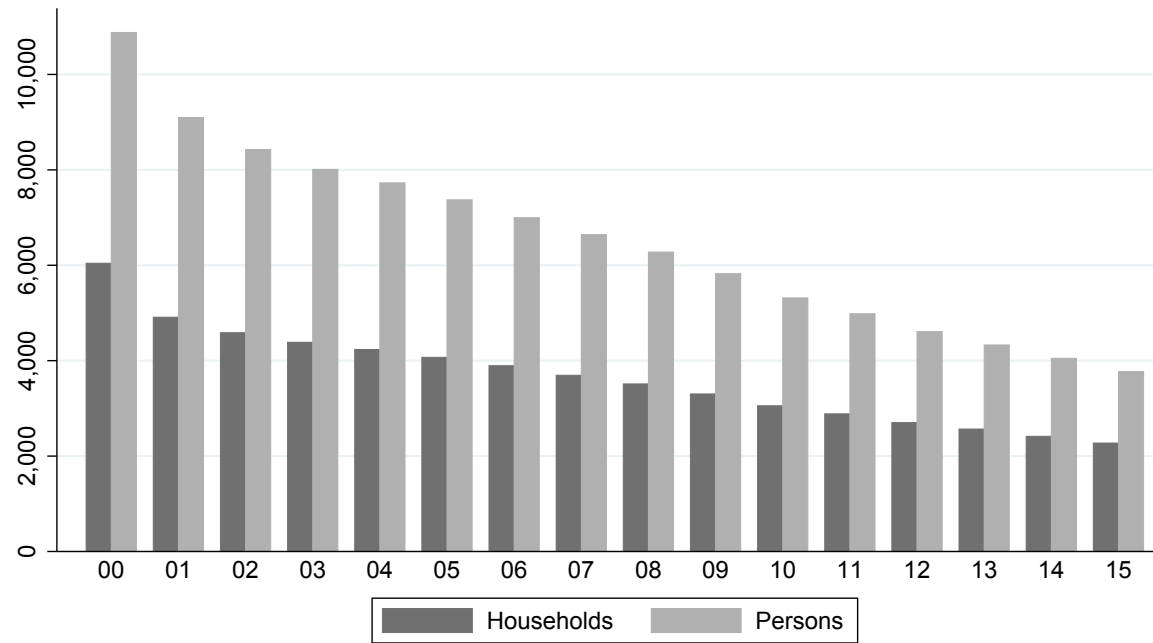


Figure 6: Comparison of Successful Interviews with Persons and Households (Subsamples F), Waves 1 to 16

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Persons	10.880	9.098	8.427	8.010	7.727	7.372	6.997	6.642	6.276	5.824	5.316	4.984	4.610	4.329	4.049	3.773
Households	6.043	4.911	4.586	4.386	4.235	4.070	3.895	3.694	3.513	3.303	3.055	2.885	2.702	2.567	2.414	2.273

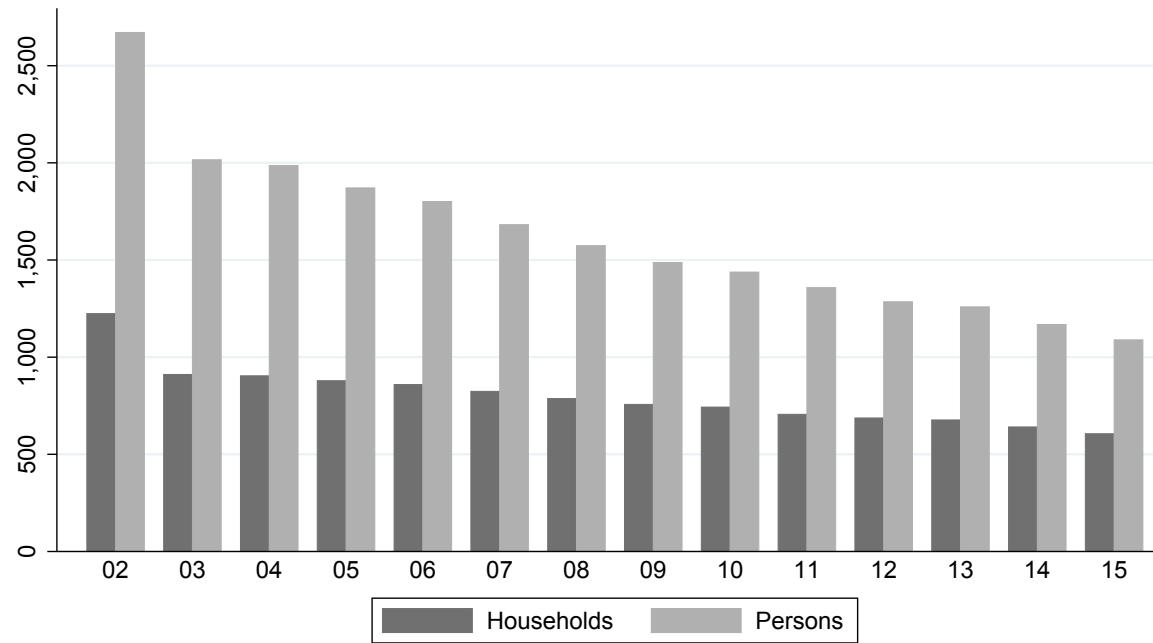


Figure 7: Comparison of Successful Interviews with Persons and Households (Subsamples G), Waves 1 to 14

Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Persons	2.671	2.016	1.986	1.871	1.801	1.682	1.574	1.487	1.438	1.358	1.285	1.259	1.168	1.089
Households	1.224	911	904	879	859	824	787	757	743	706	687	677	641	606

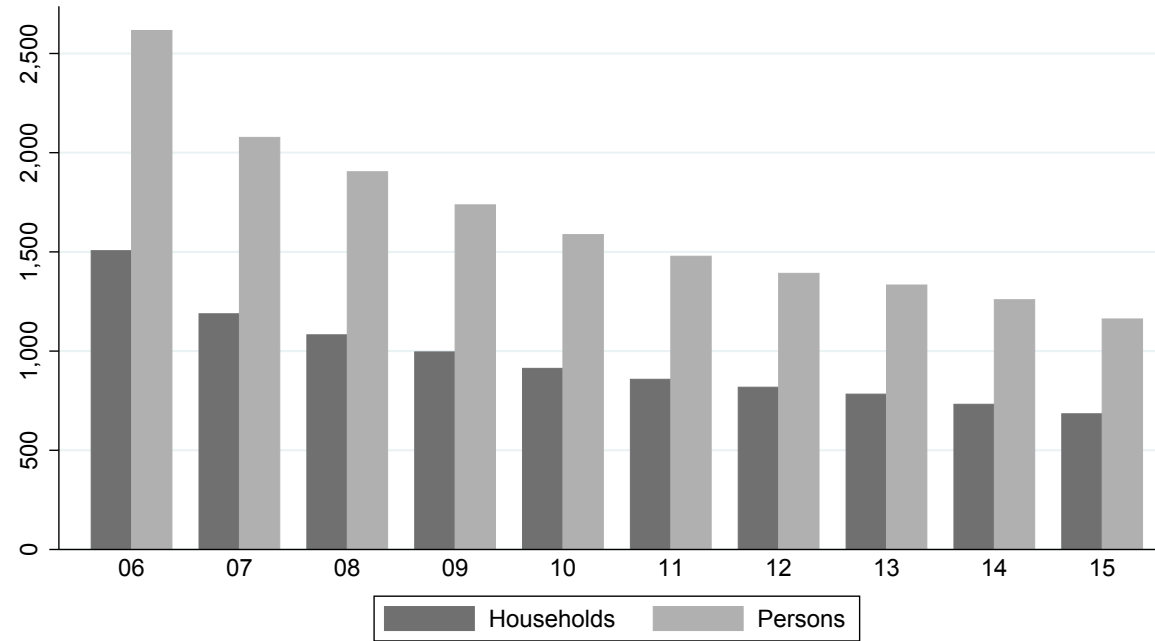


Figure 8: Comparison of Successful Interviews with Persons and Households (Subsamples H), Waves 1 to 10

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Persons	2.616	2.077	1.904	1.737	1.587	1.478	1.392	1.333	1.259	1.162
Households	1.506	1.188	1.082	996	913	858	818	783	732	684

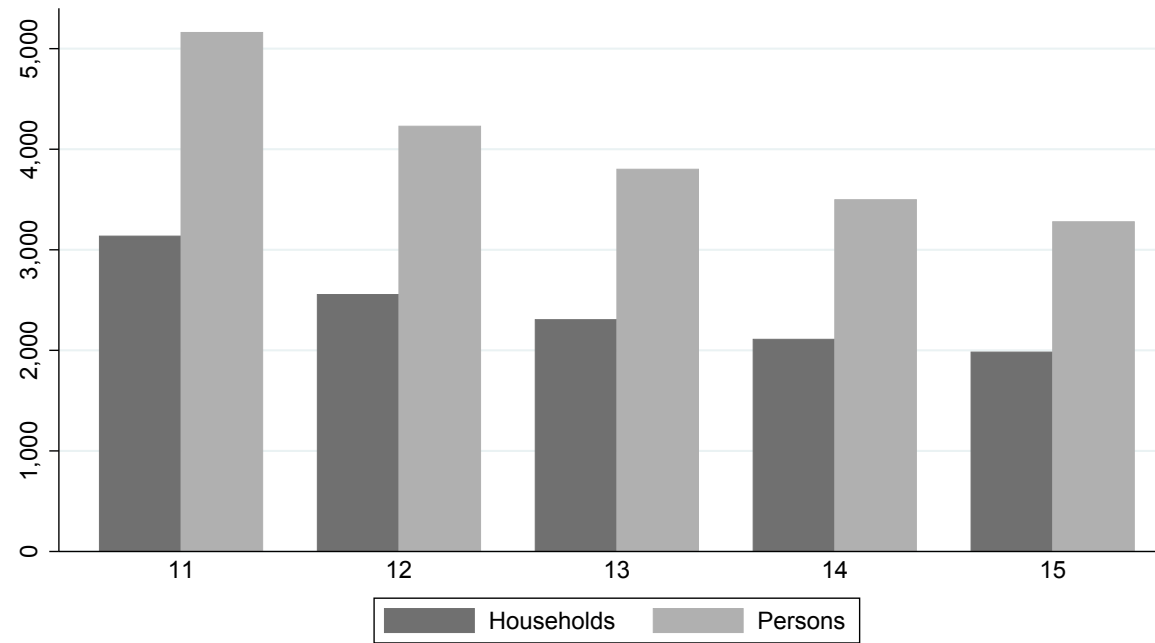


Figure 9: Comparison of Successful Interviews with Persons and Households (Subsample J), Waves 1 to 5

Year	2011	2012	2013	2014	2015
Persons	5.161	4.229	3.801	3.498	3.279
Households	3.136	2.555	2.305	2.110	1.983

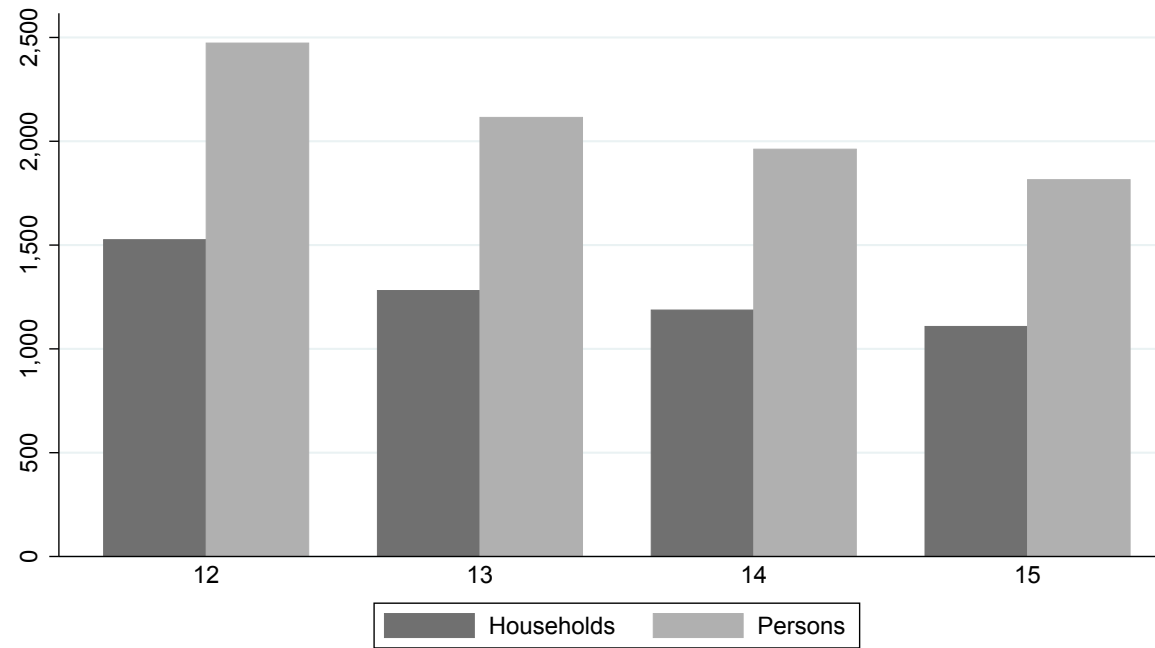


Figure 10: Comparison of Successful Interviews with Persons and Households (Subsample K), Waves 1 to 4

Year	2012	2013	2014	2015
Persons	2.473	2.115	1.962	1.815
Households	1.256	1.281	1.187	1.108

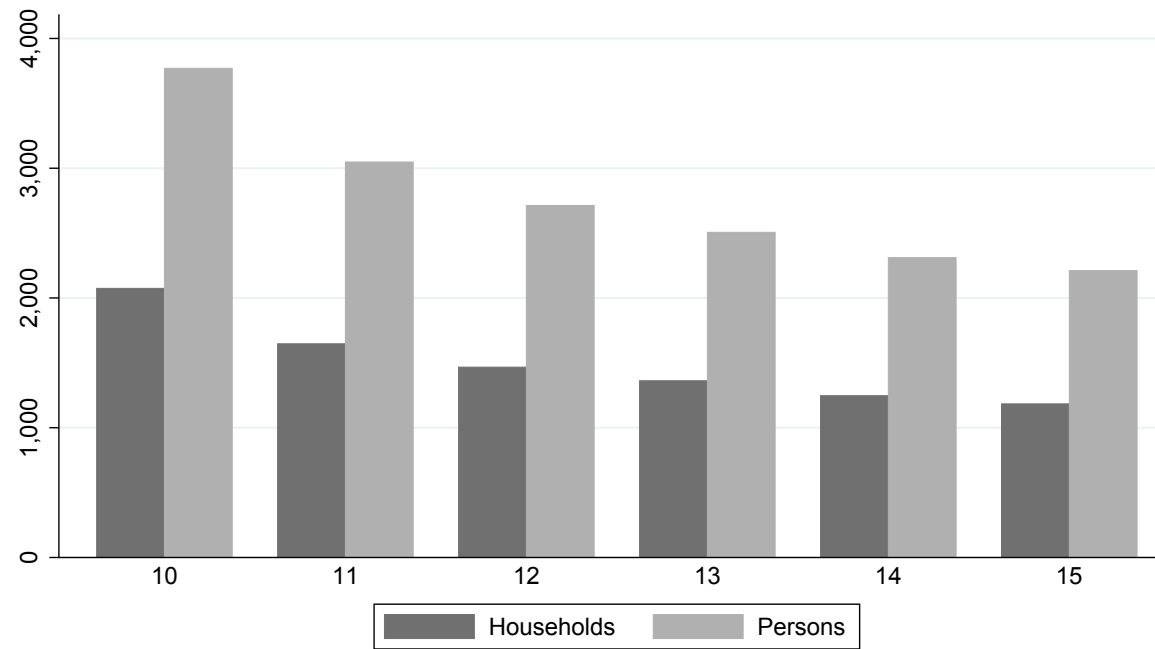


Figure 11: Comparison of Successful Interviews with Persons and Households (Subsample L1), Waves 1 to 6

Year	2010	2011	2012	2013	2014	2015
Persons	3.770	3.048	2.713	2.506	2.311	2.211
Households	2.074	1.647	1.467	1.362	1.247	1.184

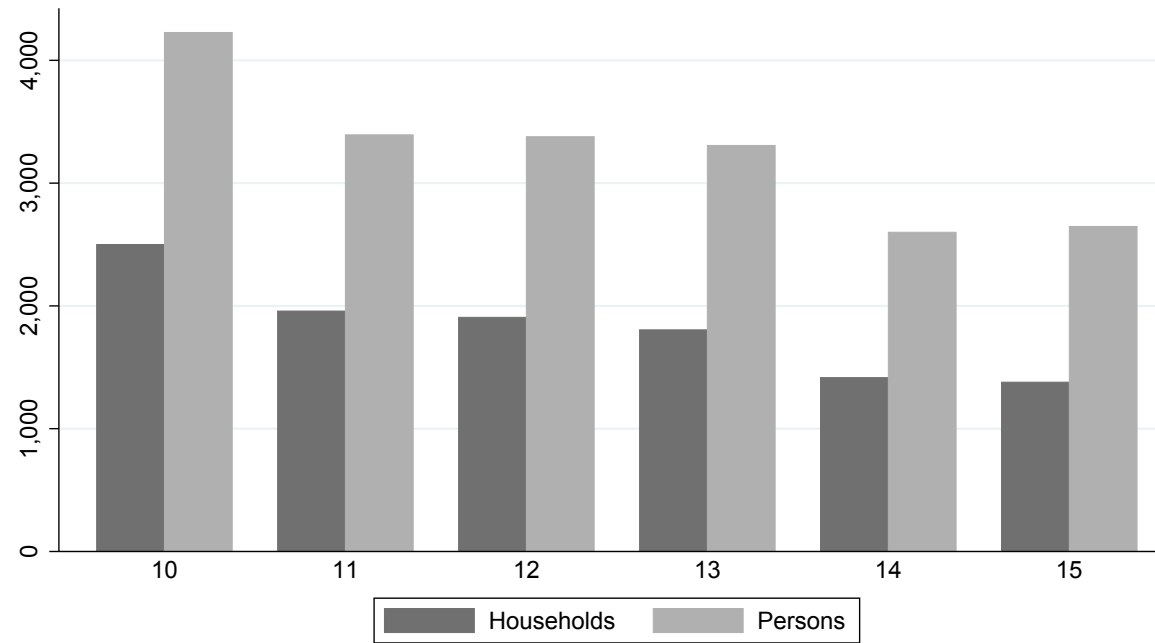


Figure 12: Comparison of Successful Interviews with Persons and Households (Subsample L2), Waves 1 to 6

Year	2010	2011	2012	2013	2014	2015
Persons	4.227	3.393	3.378	3.307	2.600	2.647
Households	2.500	1.958	1.907	1.805	1.416	1.379

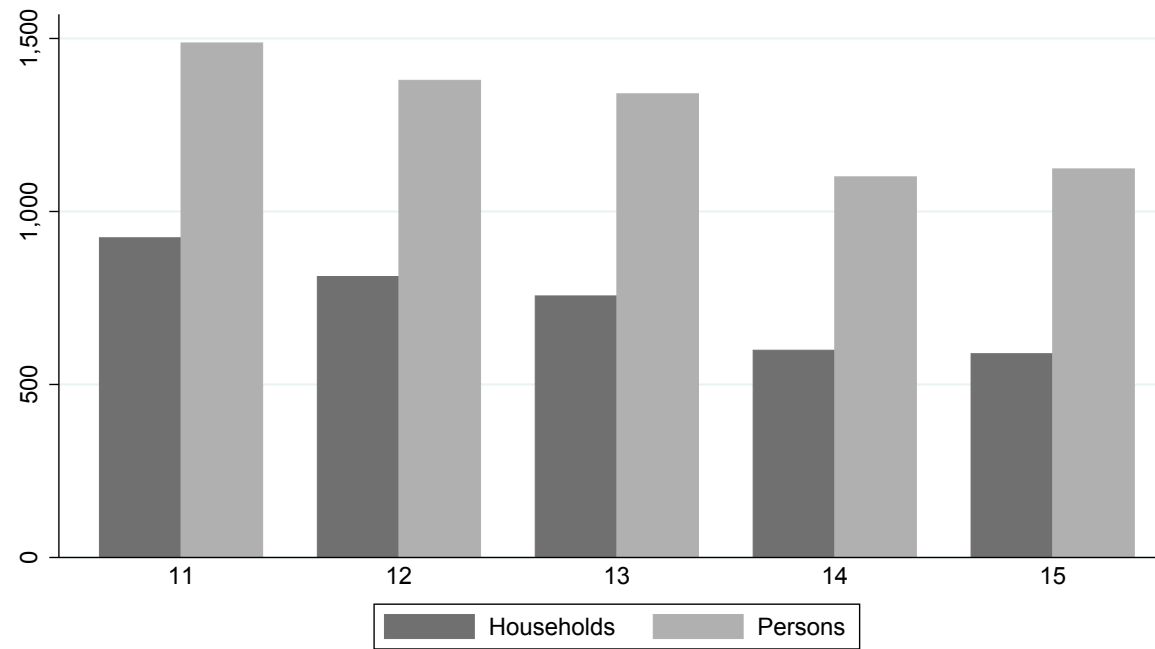


Figure 13: Comparison of Successful Interviews with Persons and Households (Subsample L3), Waves 1 to 5

Year	2011	2012	2013	2014	2015
Persons	1.487	1.379	1.340	1.100	1.123
Households	924	812	756	599	589

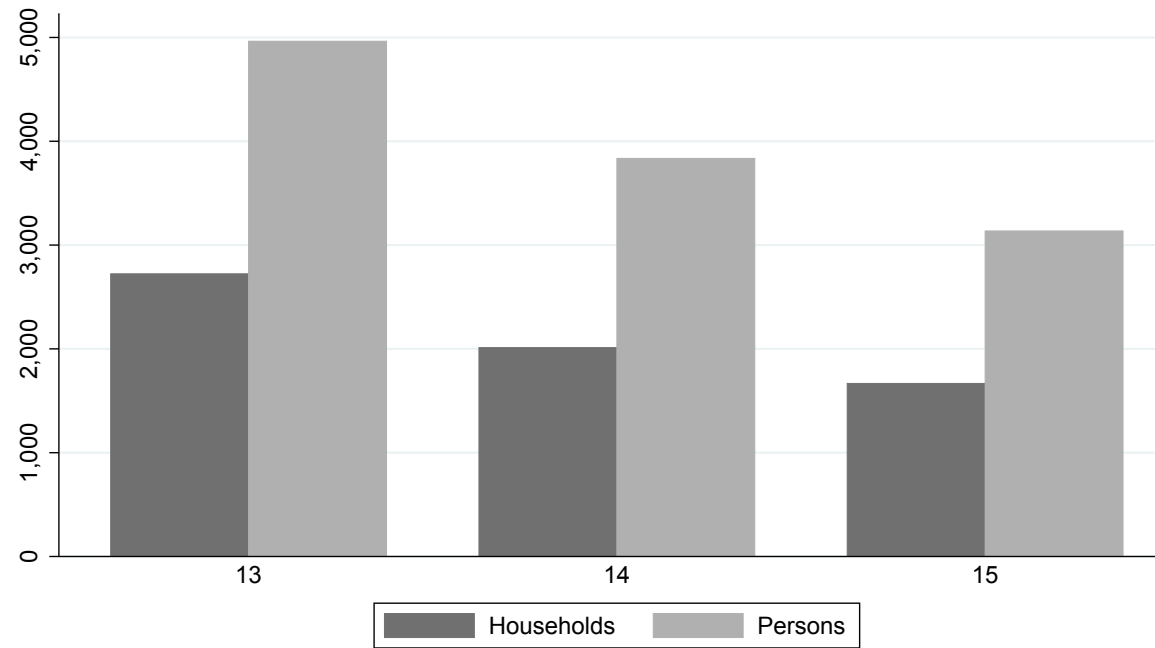


Figure 14: Comparison of Successful Interviews with Persons and Households (Subsample M1), Waves 1 to 3

Year	2013	2014	2015
Persons	4.964	3.835	3.136
Households	2.723	2.012	1.667

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 (“With interview”), exits due to survey-unrelated attrition (“Moved abroad”, “Deceased”, “Under the age of 16”), and exits due to survey-related attrition (“Temporary drop-out”, “Drop-out”).

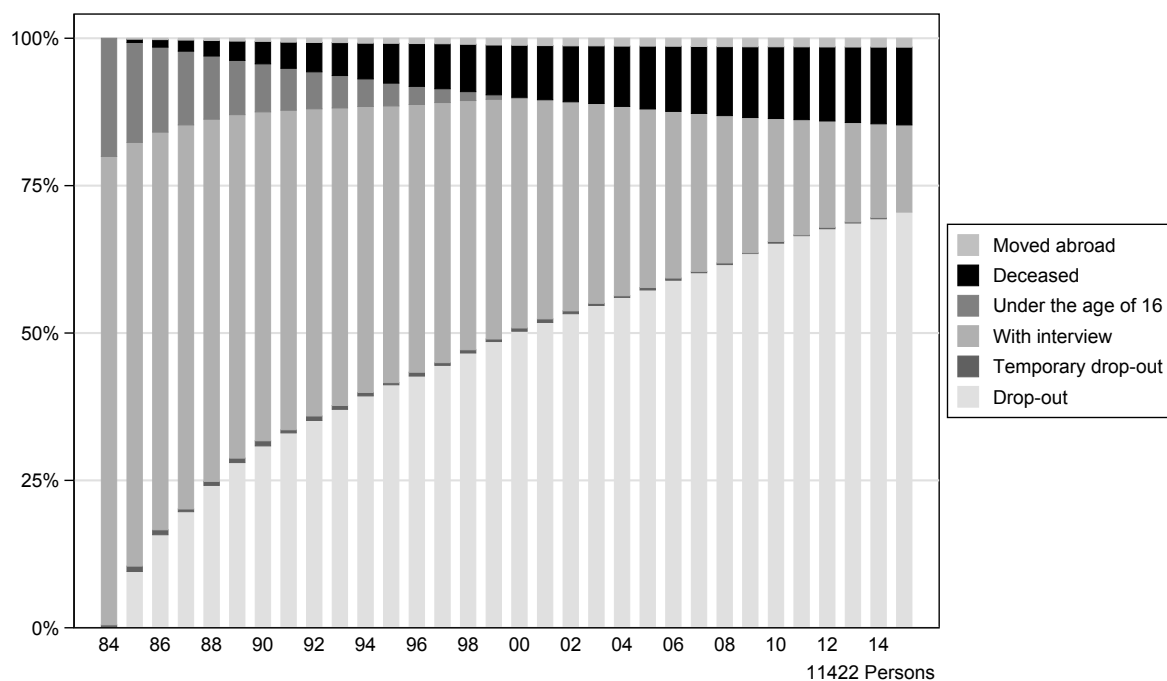


Figure 15: All First-Wave Persons (Gross Subsample A). Development up to Wave 32

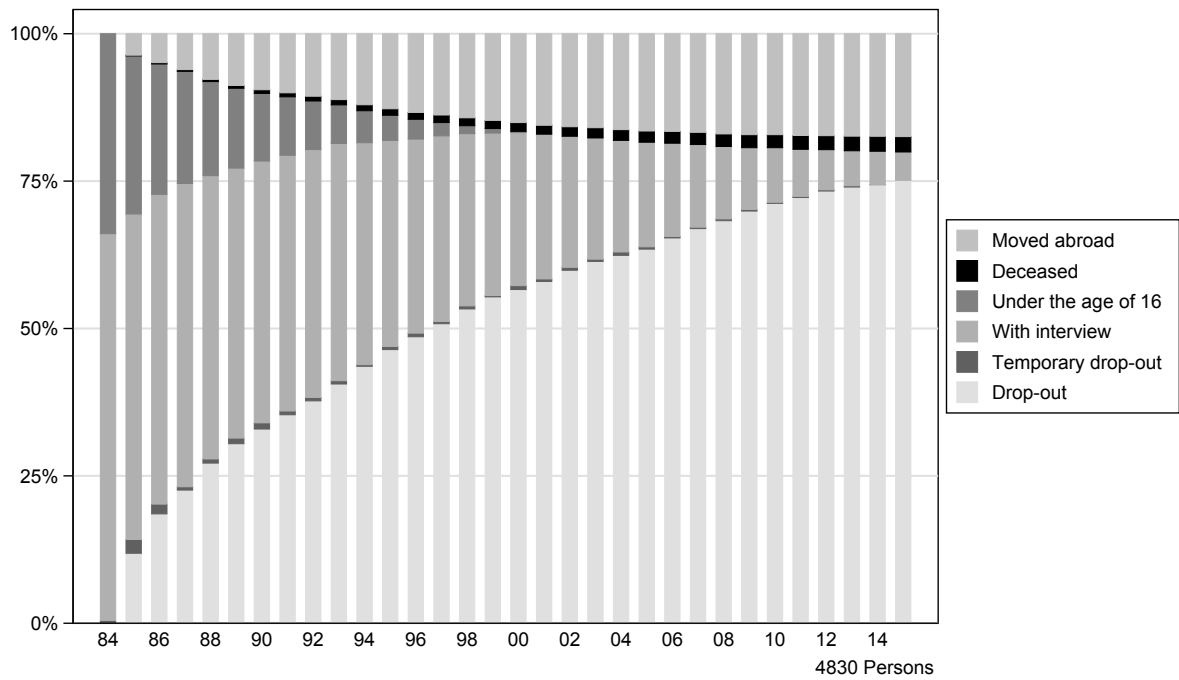


Figure 16: All First-Wave Persons (Gross Subsample B). Development up to Wave 32

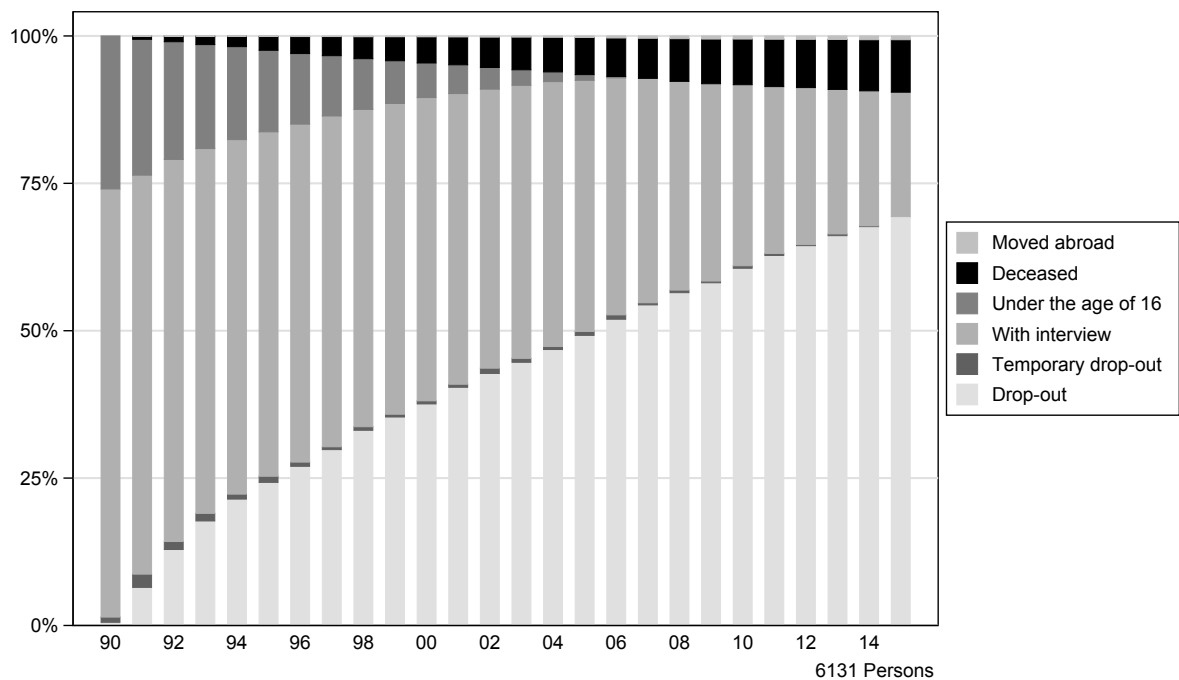


Figure 17: All First-Wave Persons (Gross Subsample C). Development up to Wave 26

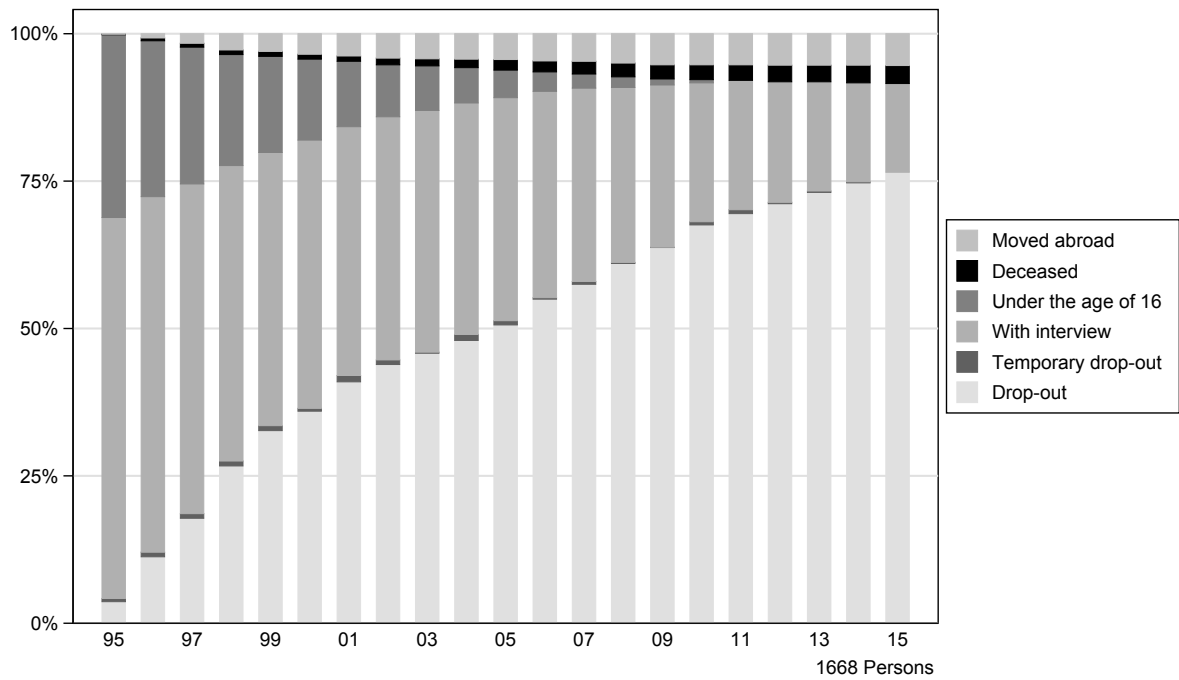


Figure 18: All First-Wave Persons (Gross Subsample D). Development up to Wave 21

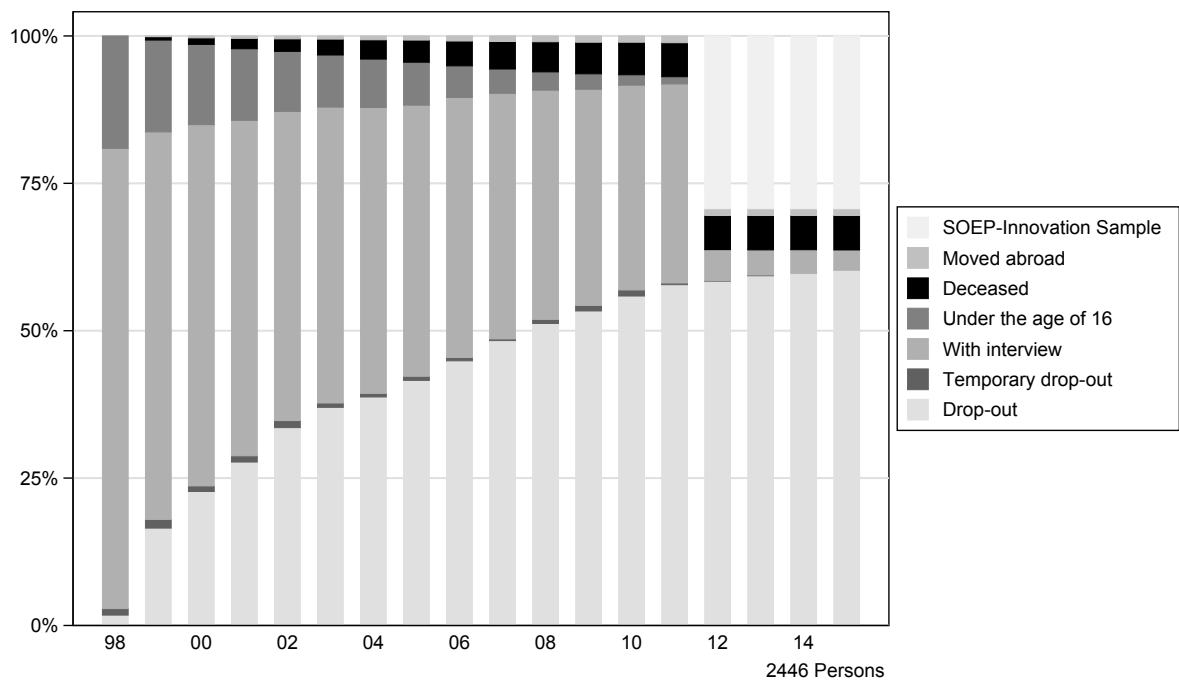


Figure 19: All First-Wave Persons (Gross Subsample E). Development up to Wave 18

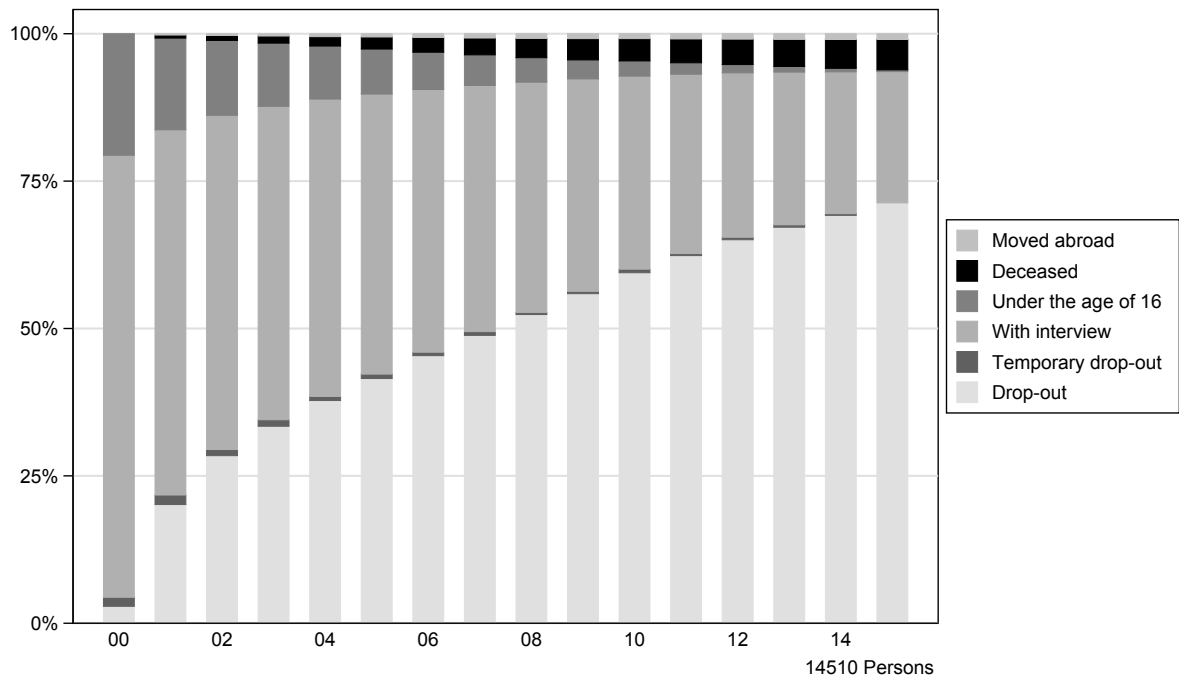


Figure 20: All First-Wave Persons (Gross Subsample F). Development up to Wave 16

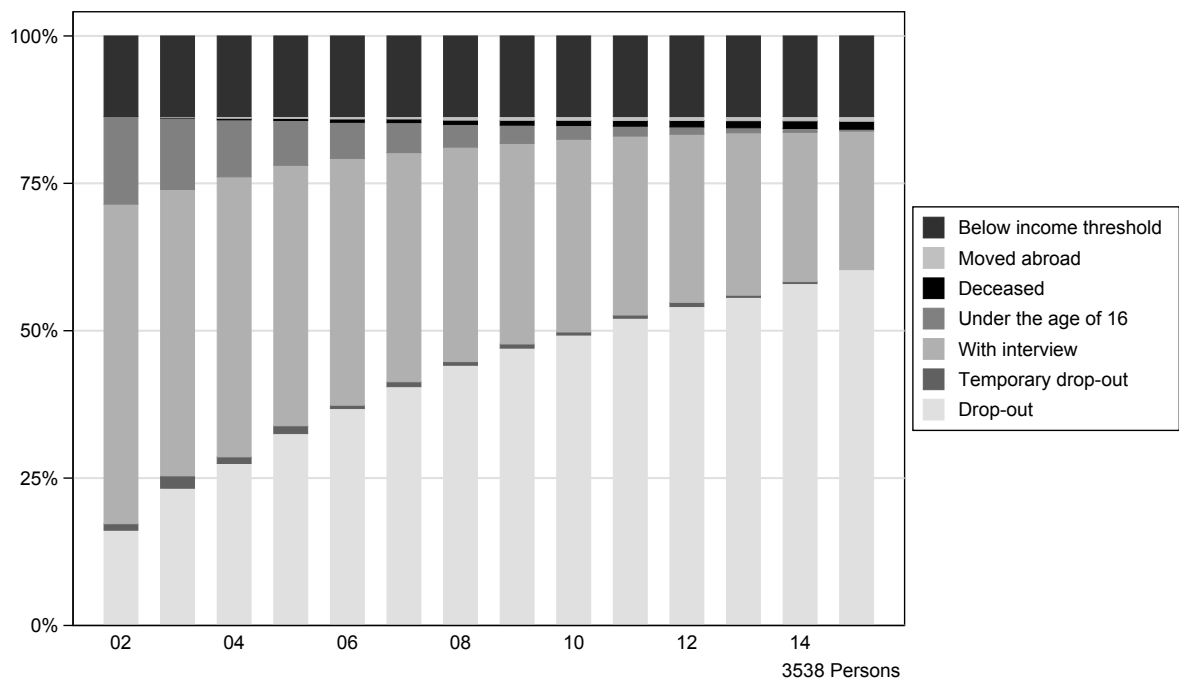


Figure 21: All First-Wave Persons (Gross Subsample G). Development up to Wave 14

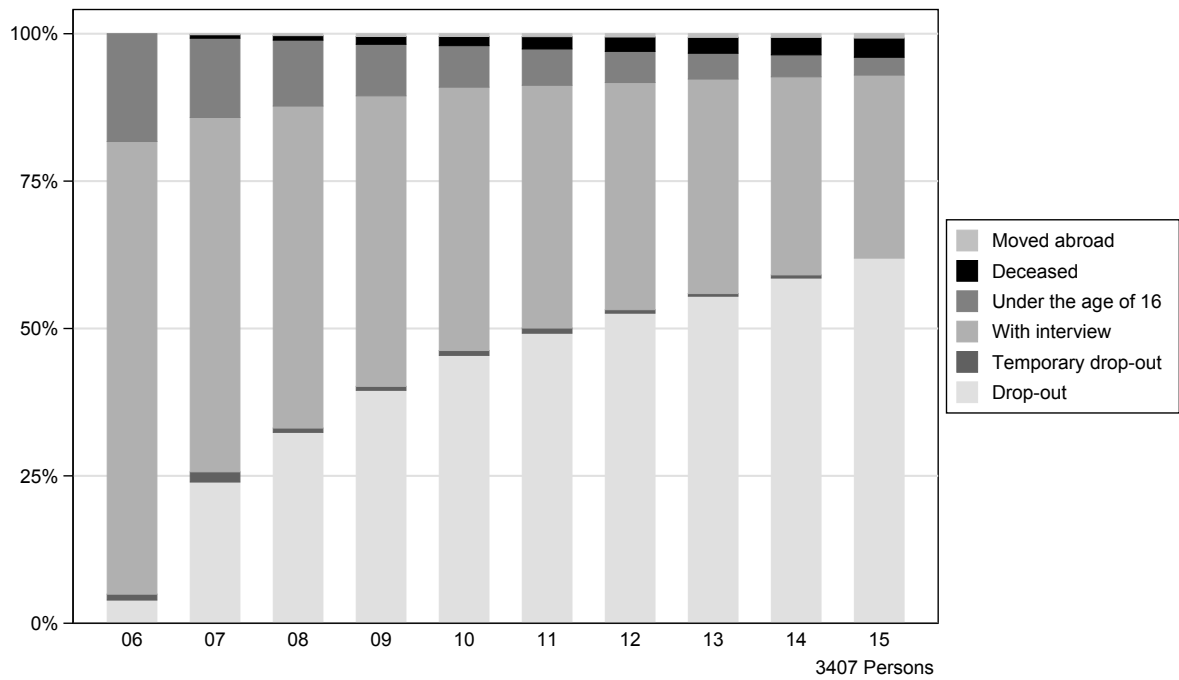


Figure 22: All First-Wave Persons (Gross Subsample H). Development up to Wave 10

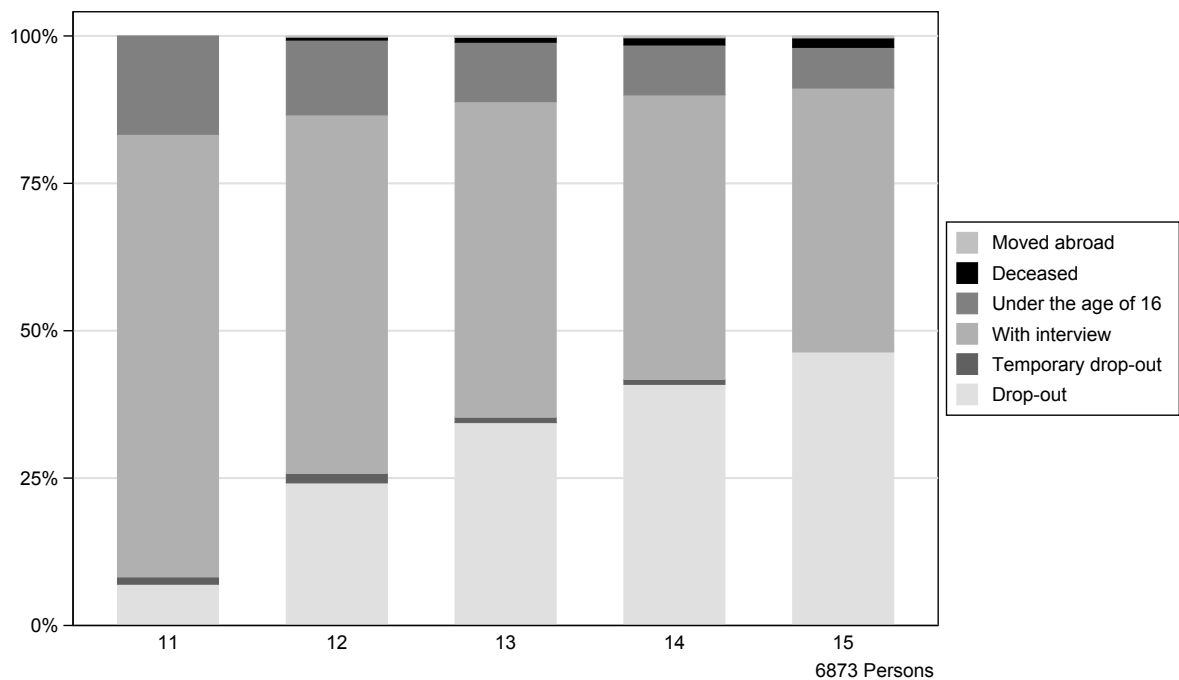


Figure 23: All First-Wave Persons (Gross Subsample J). Development up to Wave 5

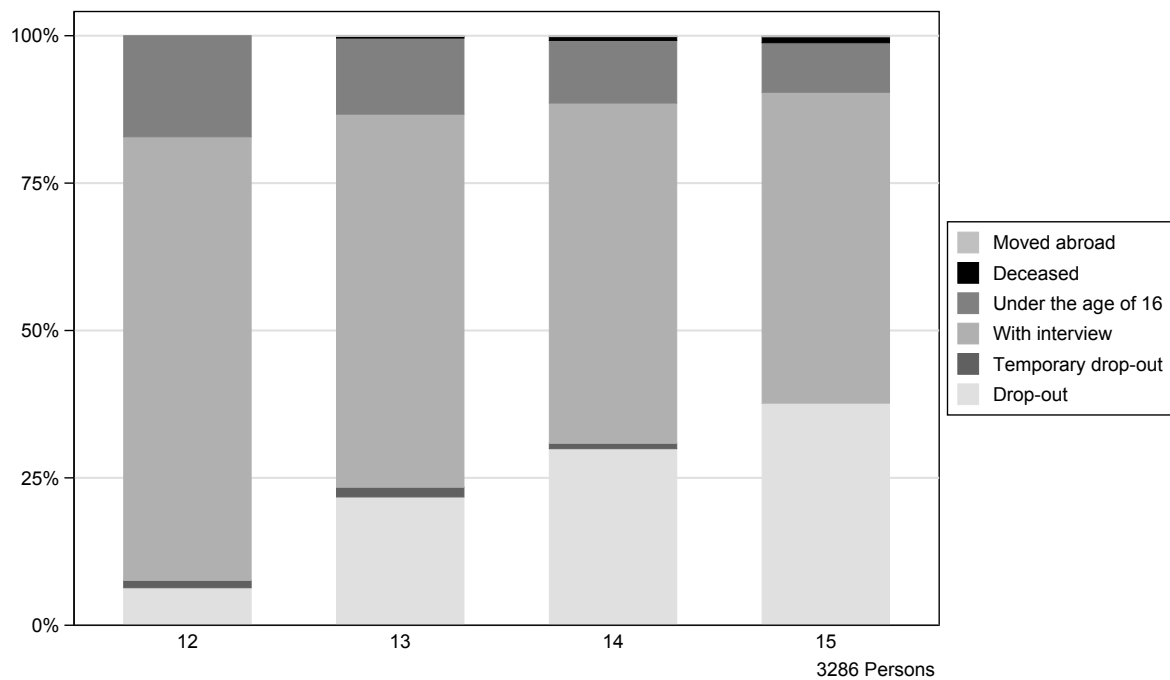


Figure 24: All First-Wave Persons (Gross Subsample K). Development up to Wave 4

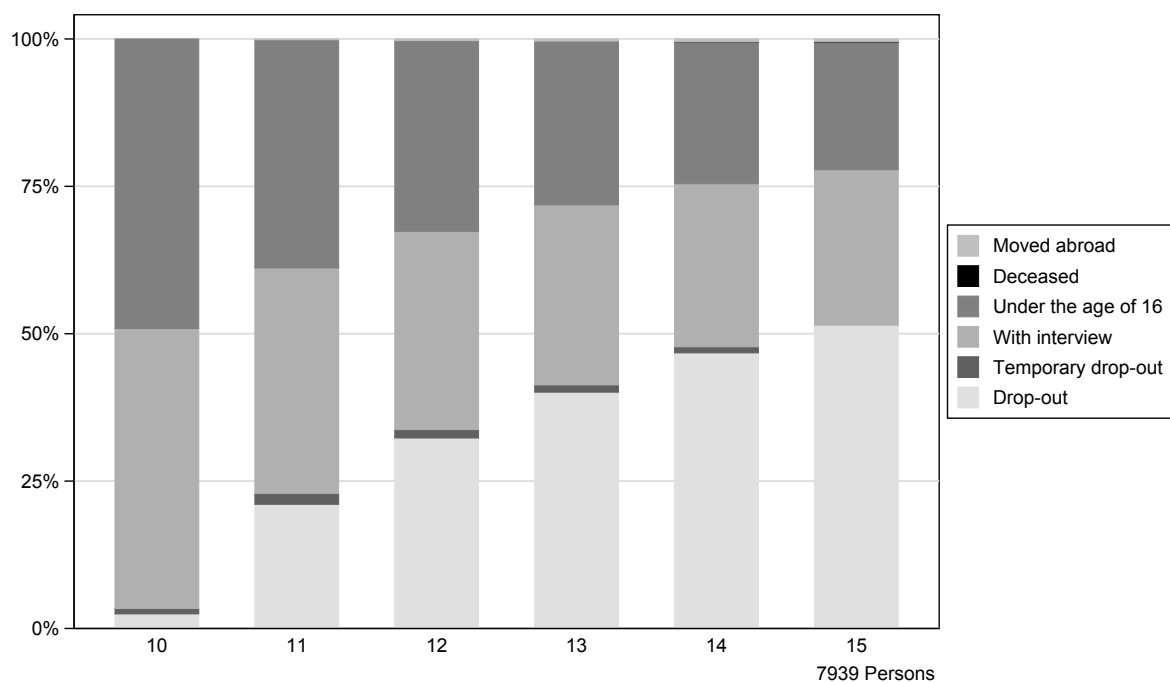


Figure 25: All First-Wave Persons (Gross Subsample L1). Development up to Wave 6

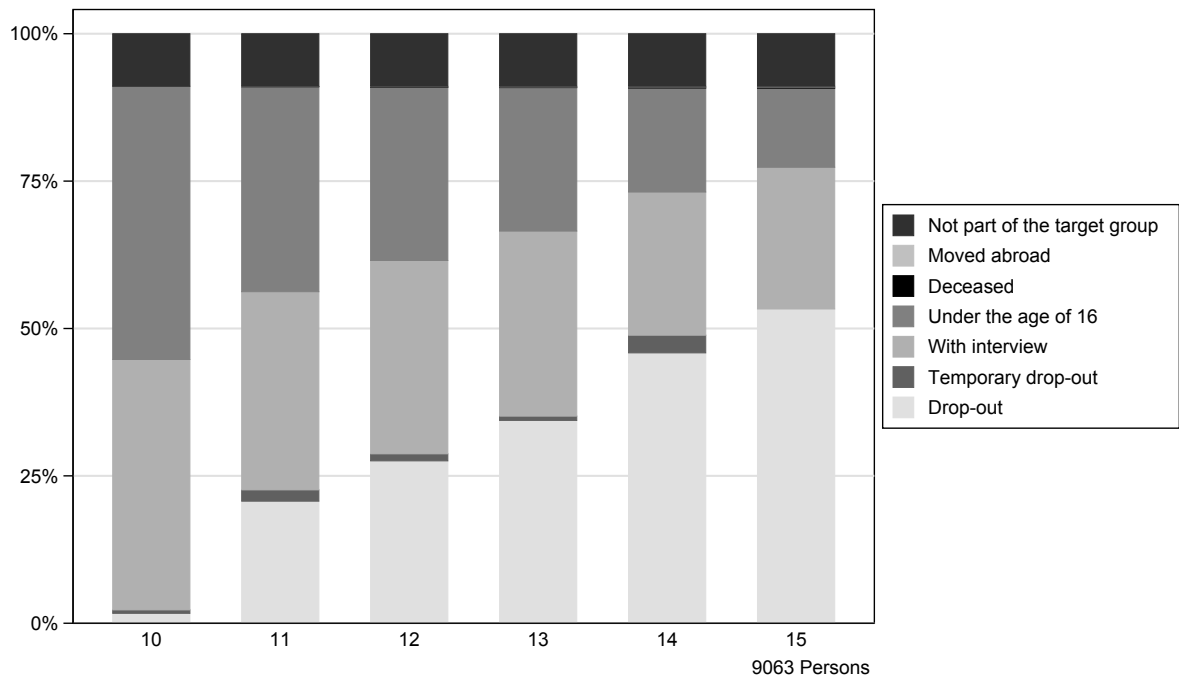


Figure 26: All First-Wave Persons (Gross Subsample L2). Development up to Wave 6

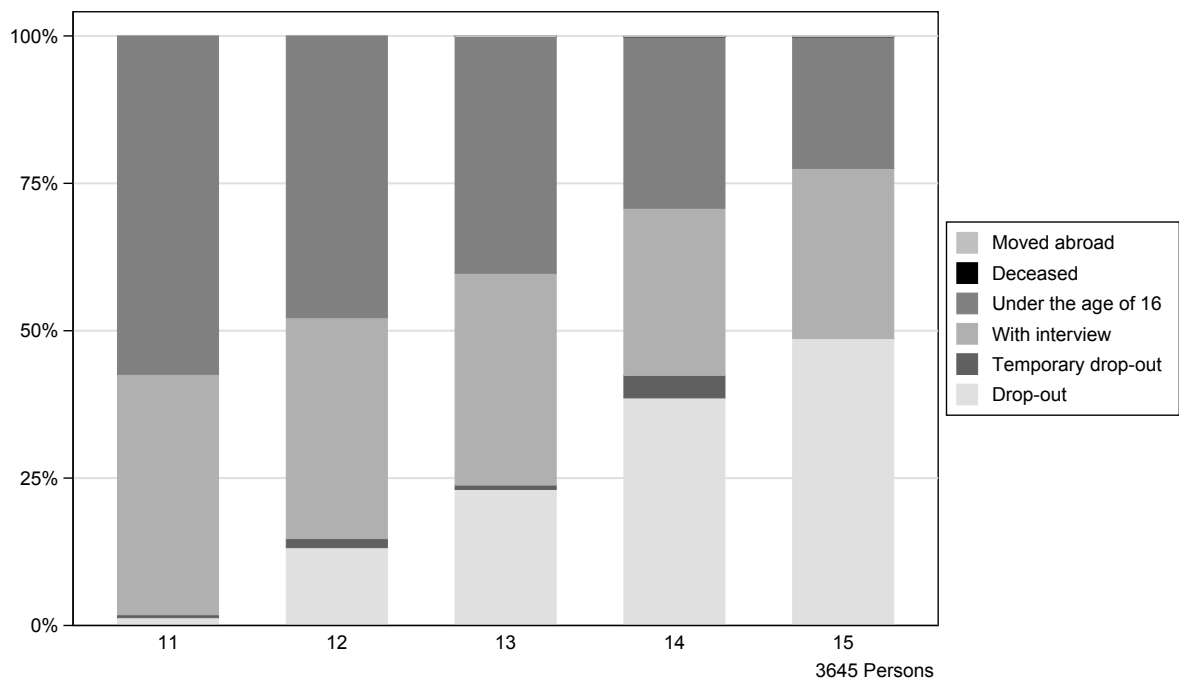


Figure 27: All First-Wave Persons (Gross Subsample L3). Development up to Wave 5

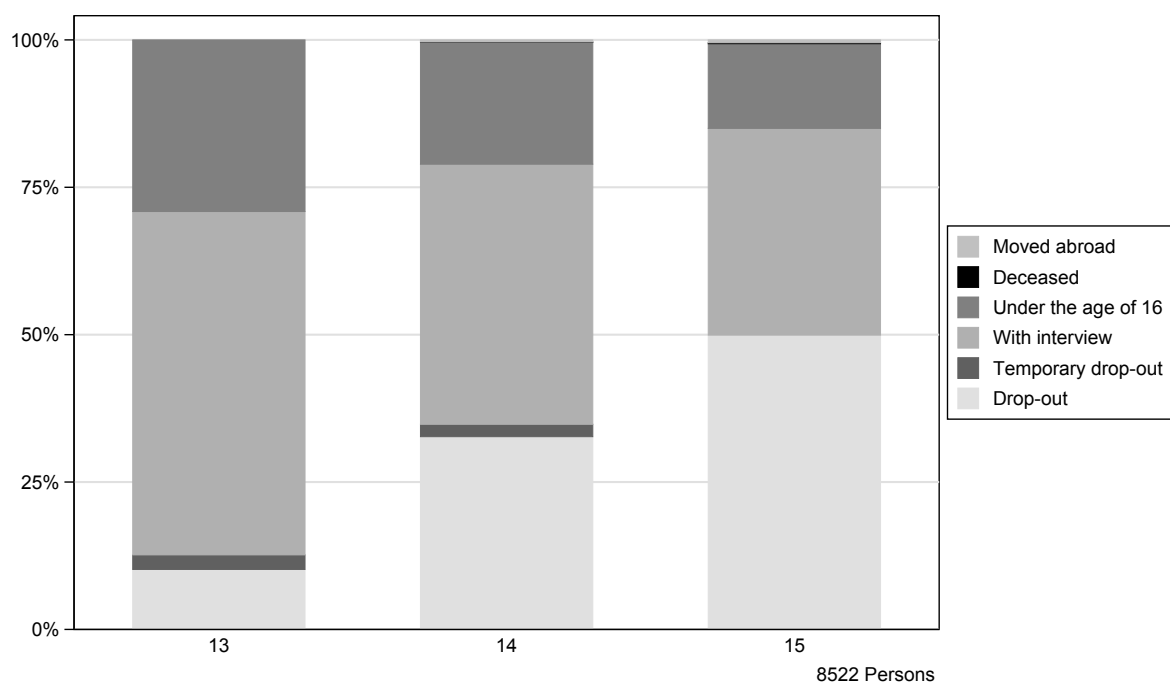


Figure 28: All First-Wave Persons (Gross Subsample M1). Development up to Wave 3

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.

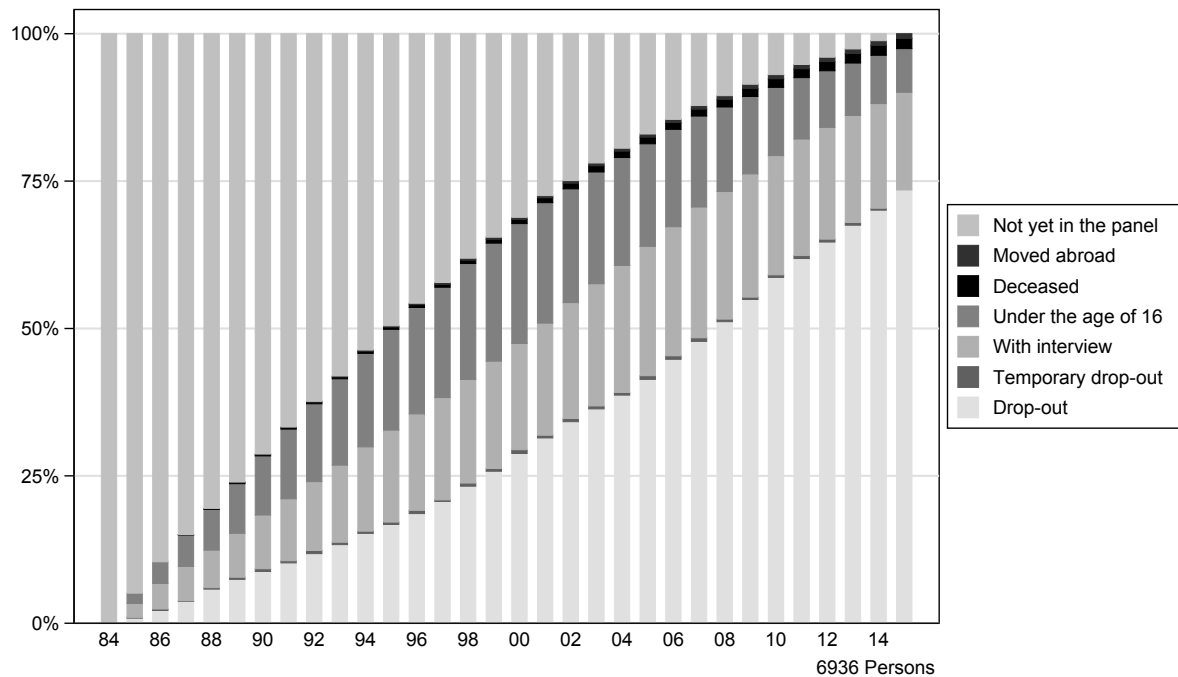


Figure 29: Entrants and their Participation Behavior (Subsample A). Development up to Wave 32

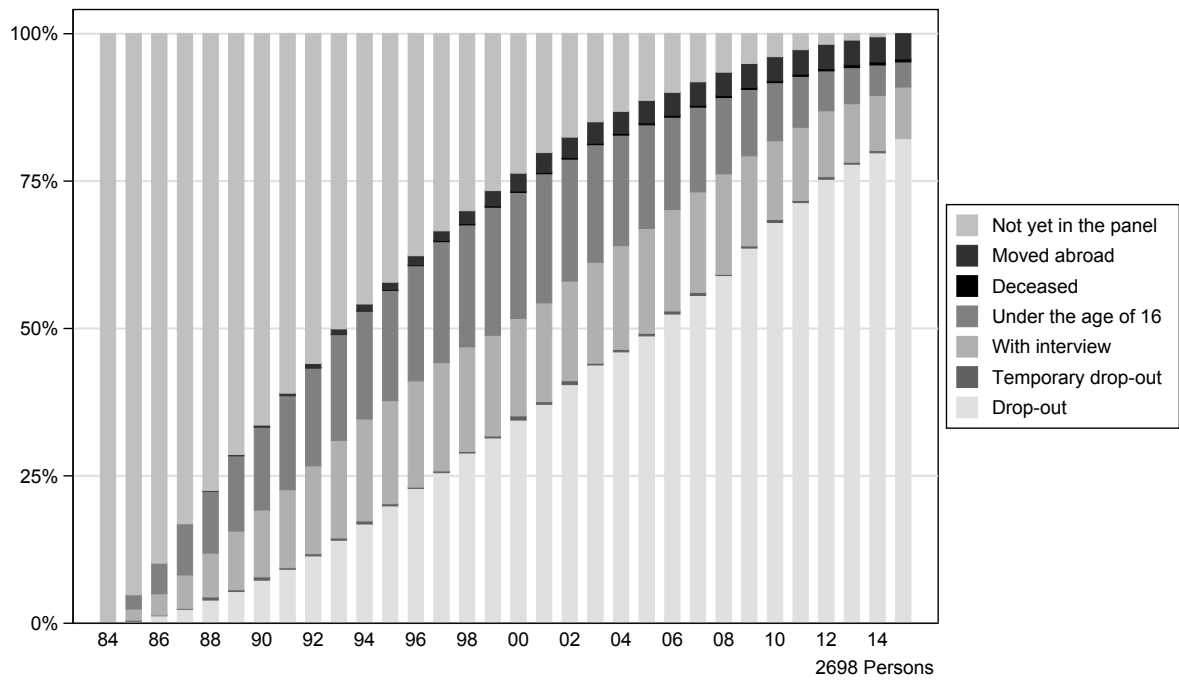


Figure 30: Entrants and their Participation Behavior (Subsample B). Development up to Wave 32

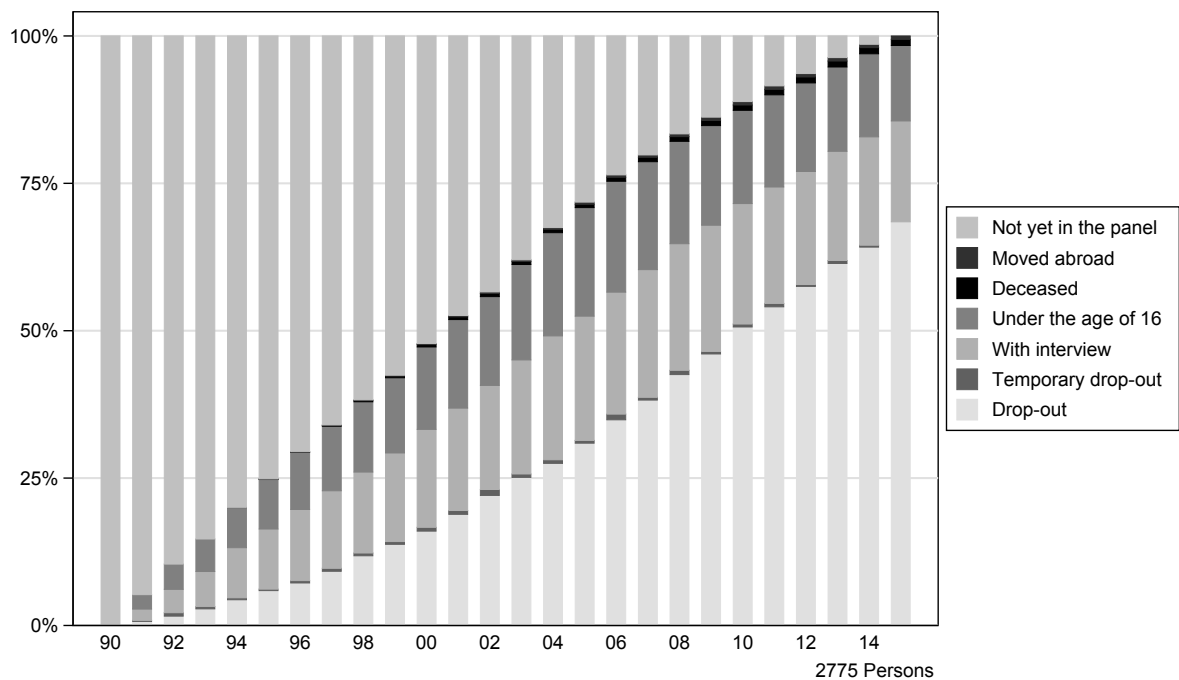


Figure 31: Entrants and their Participation Behavior (Subsample C). Development up to Wave 26

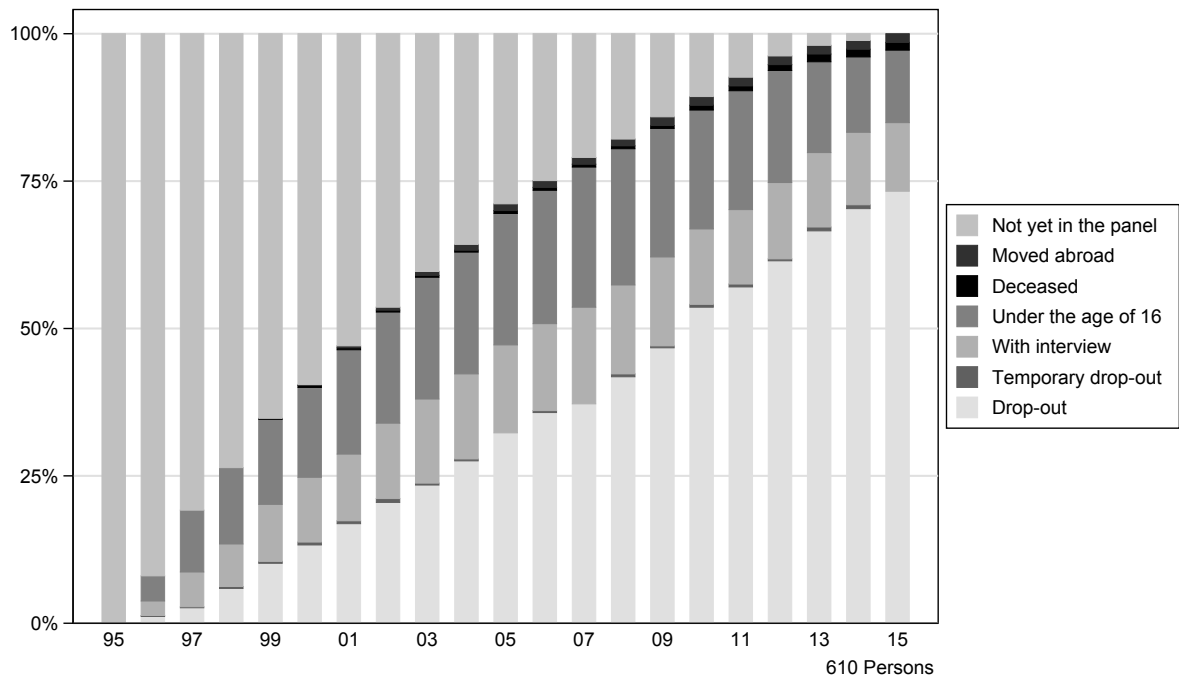


Figure 32: Entrants and their Participation Behavior (Subsample D). Development up to Wave 21

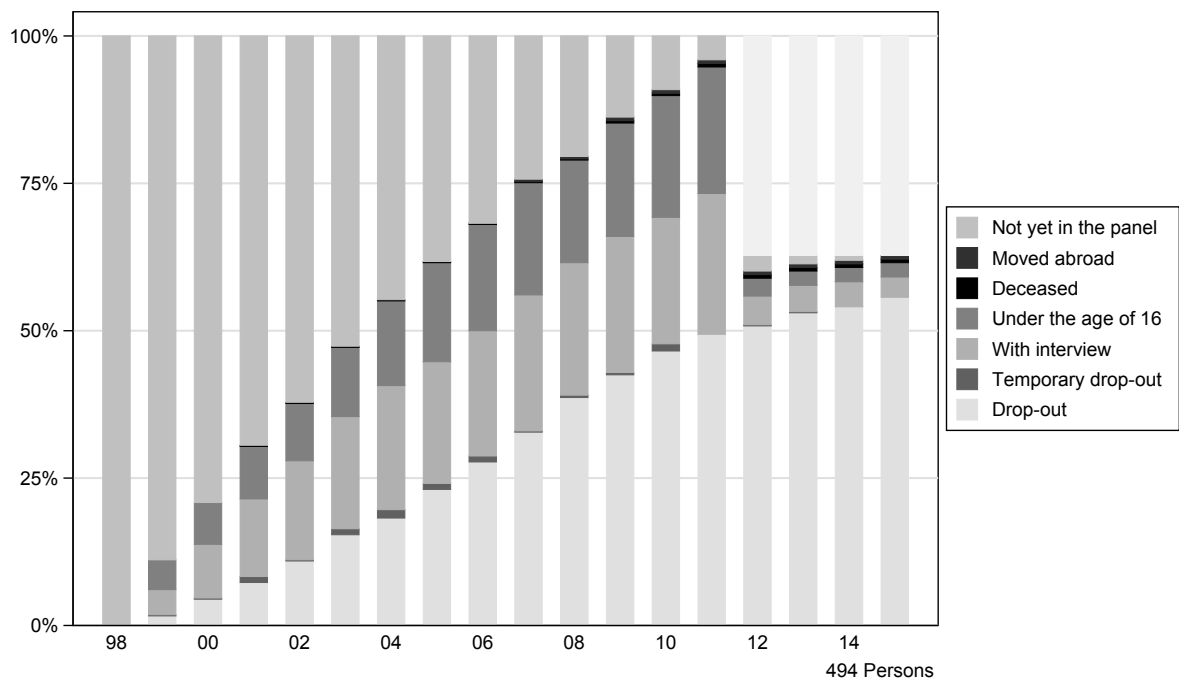


Figure 33: Entrants and their Participation Behavior (Subsample E). Development up to Wave 18

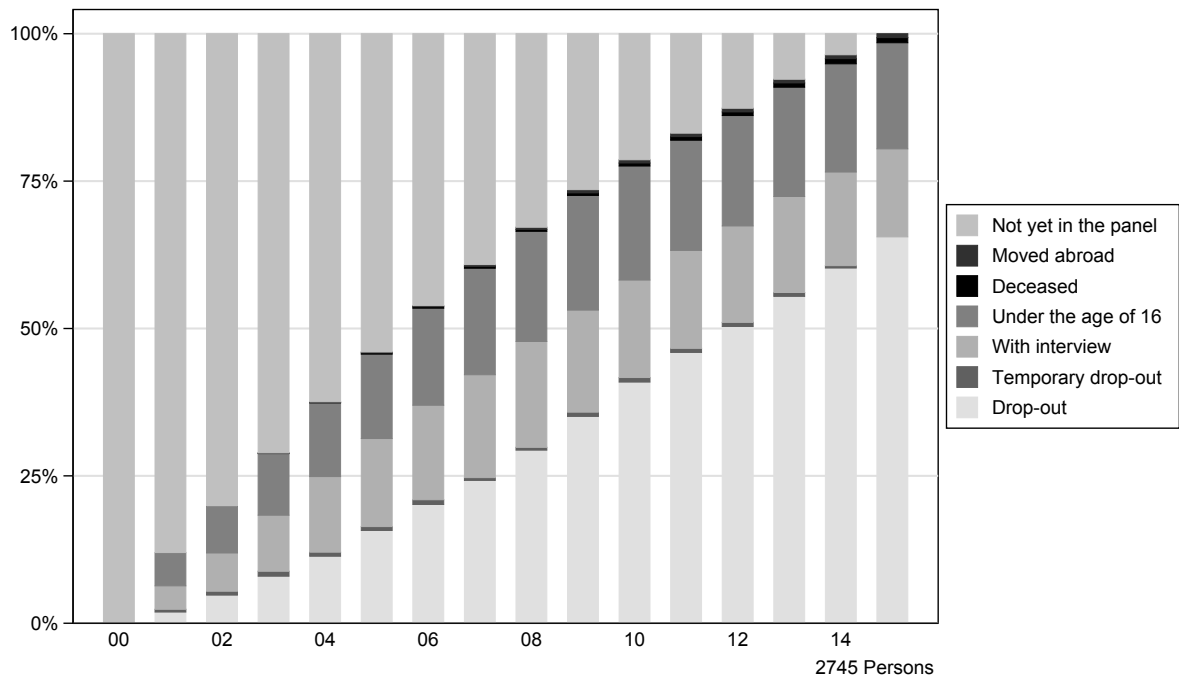


Figure 34: Entrants and their Participation Behavior (Subsample F). Development up to Wave 16

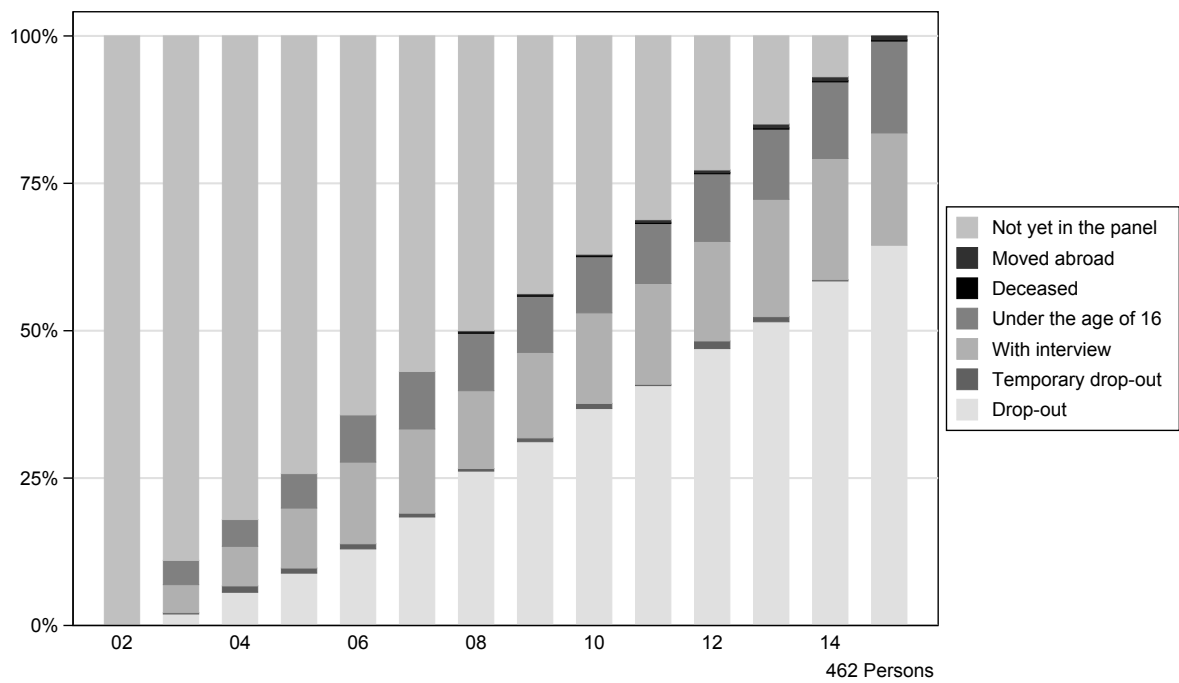


Figure 35: Entrants and their Participation Behavior (Subsample G). Development up to Wave 14

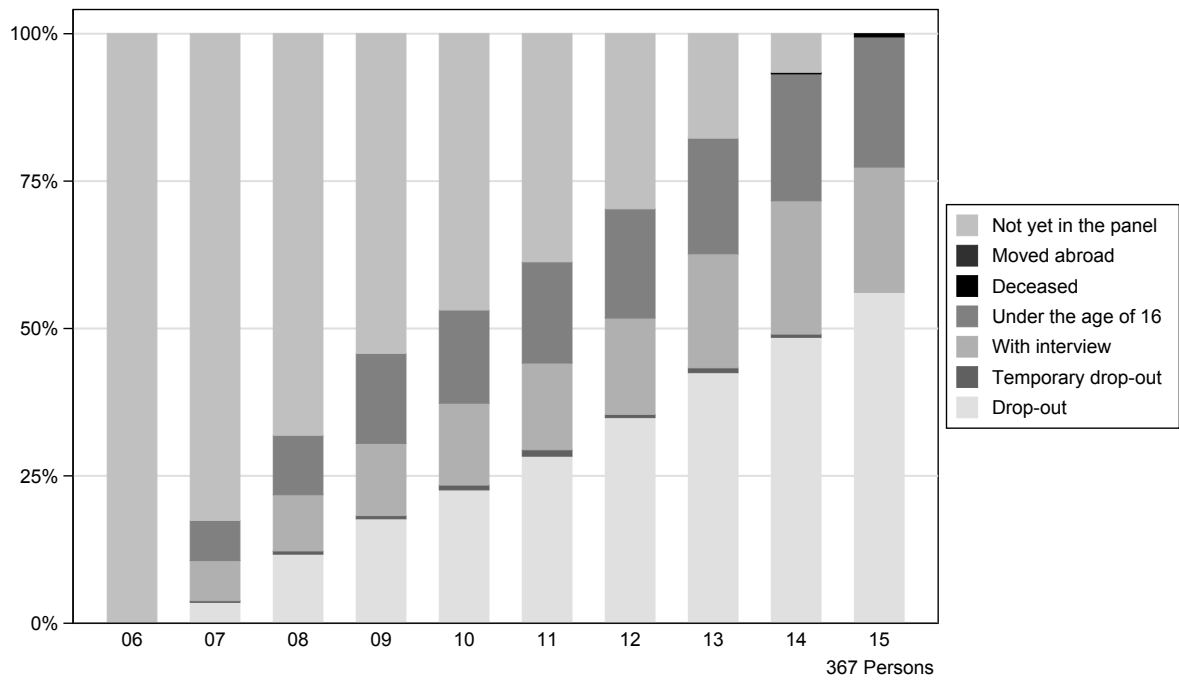


Figure 36: Entrants and their Participation Behavior (Subsample H). Development up to Wave 10

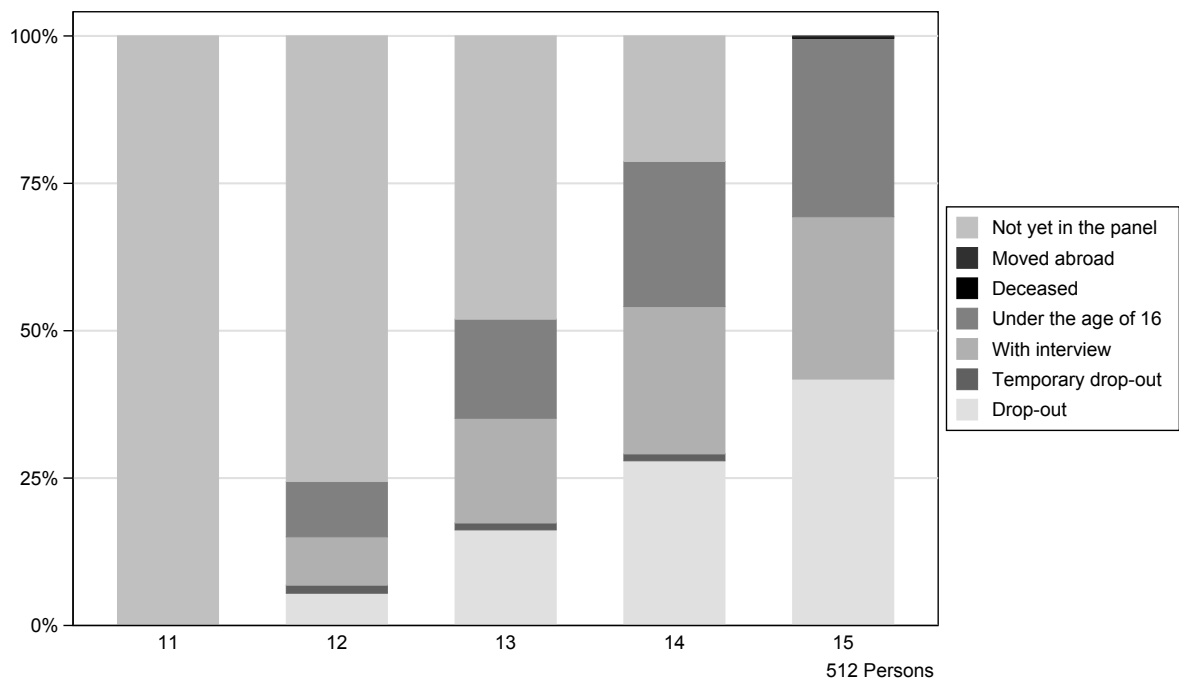


Figure 37: Entrants and their Participation Behavior (Subsample J). Development up to Wave 5

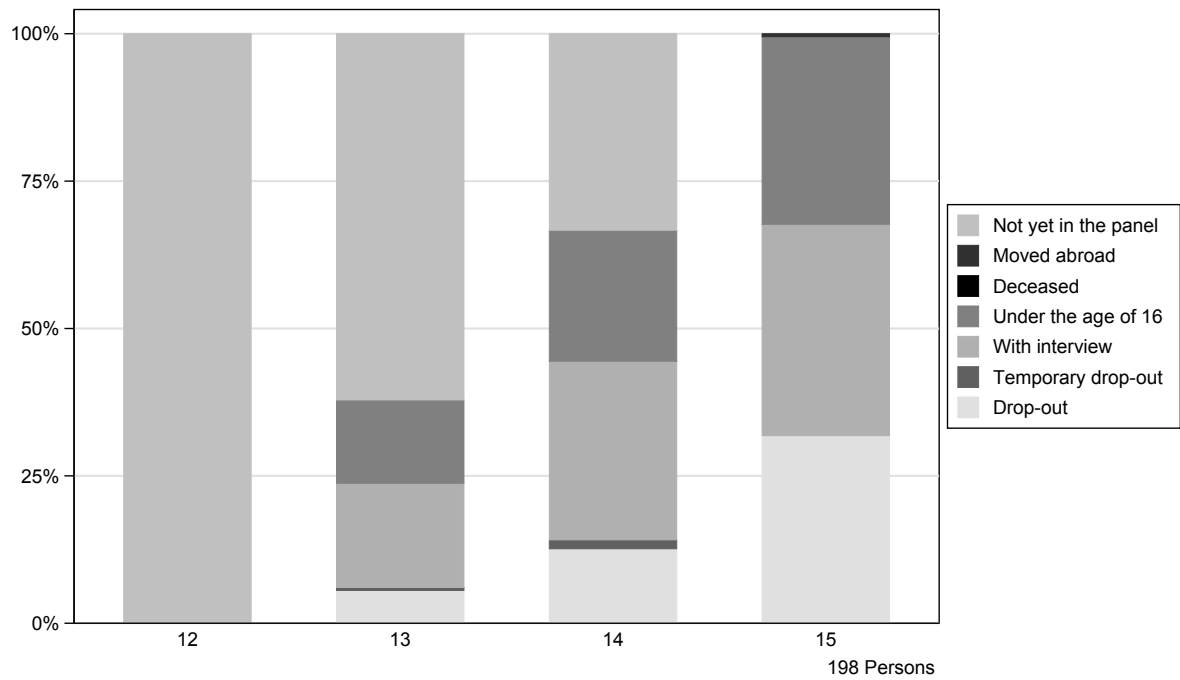


Figure 38: Entrants and their Participation Behavior (Subsample K). Development up to Wave 4

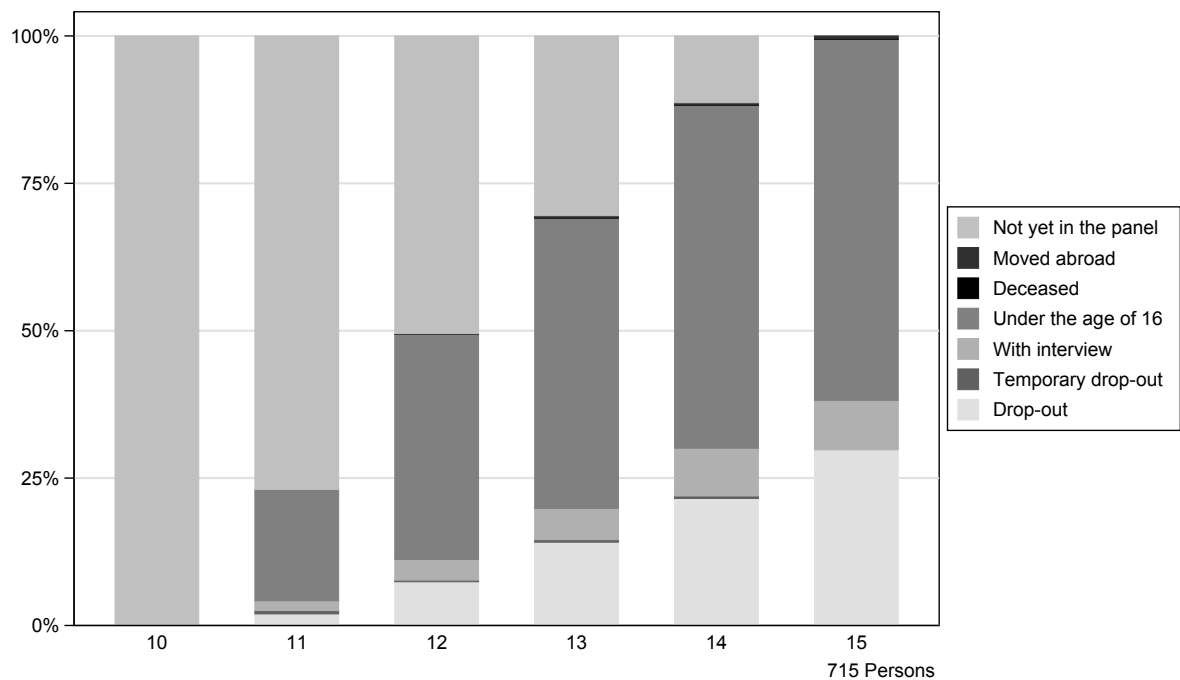


Figure 39: Entrants and their Participation Behavior (Subsample L1). Development up to Wave 6

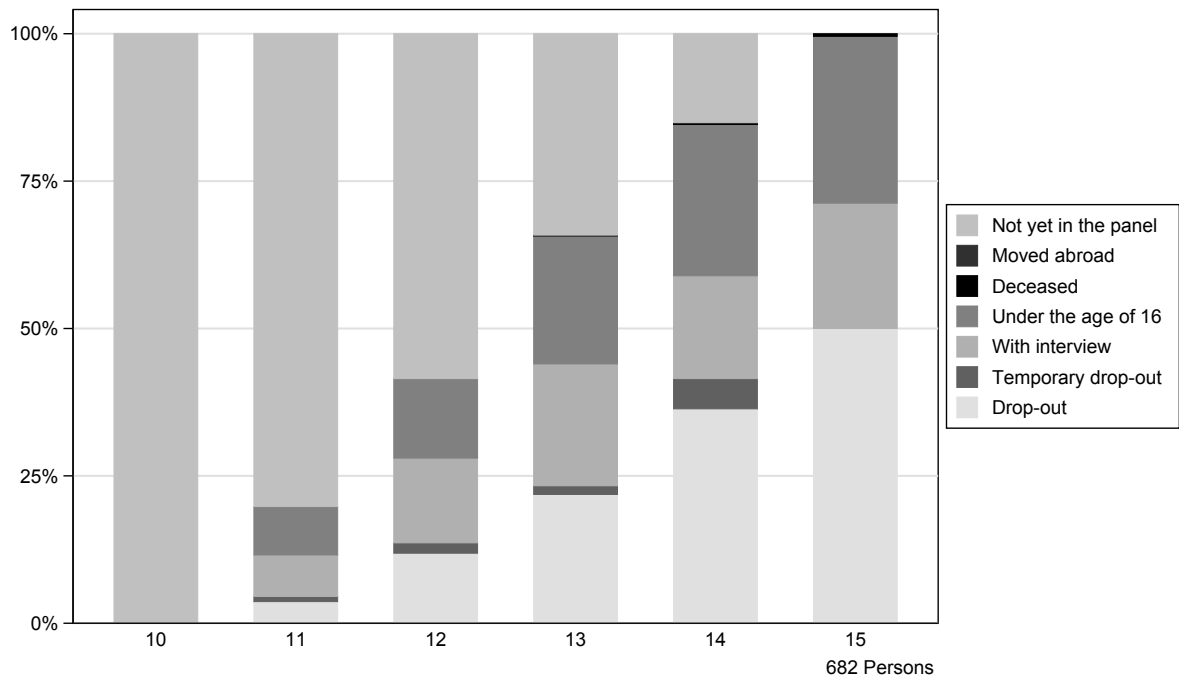


Figure 40: Entrants and their Participation Behavior (Subsample L2). Development up to Wave 6

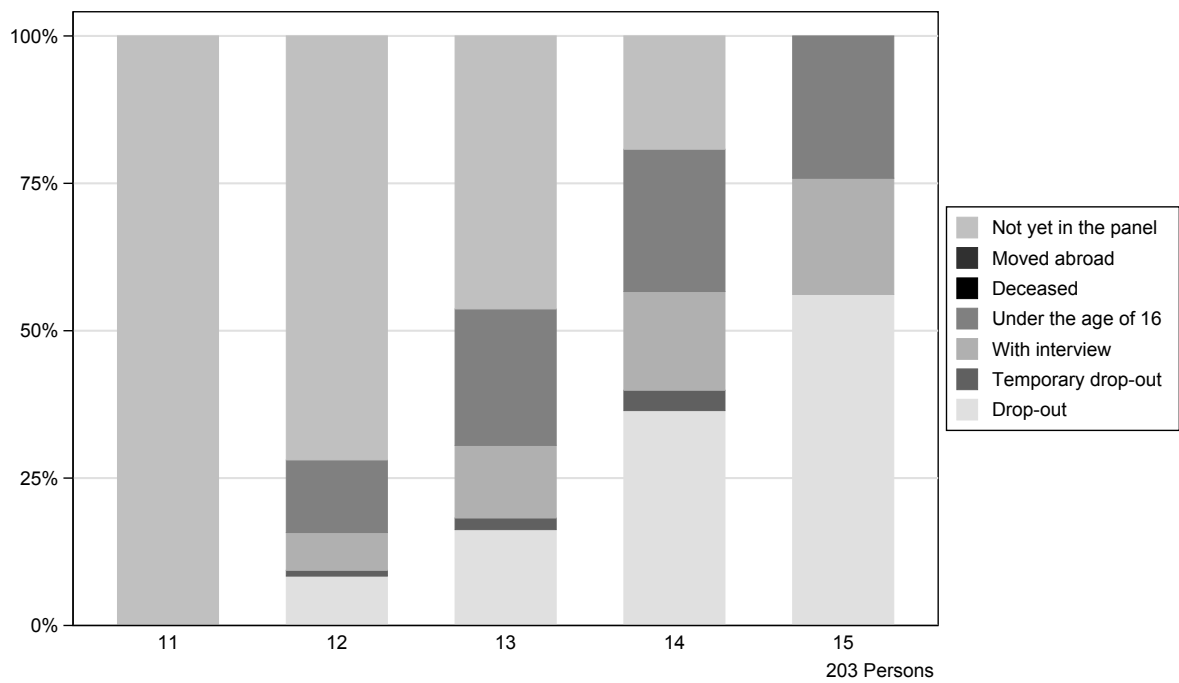


Figure 41: Entrants and their Participation Behavior (Subsample L3). Development up to Wave 5

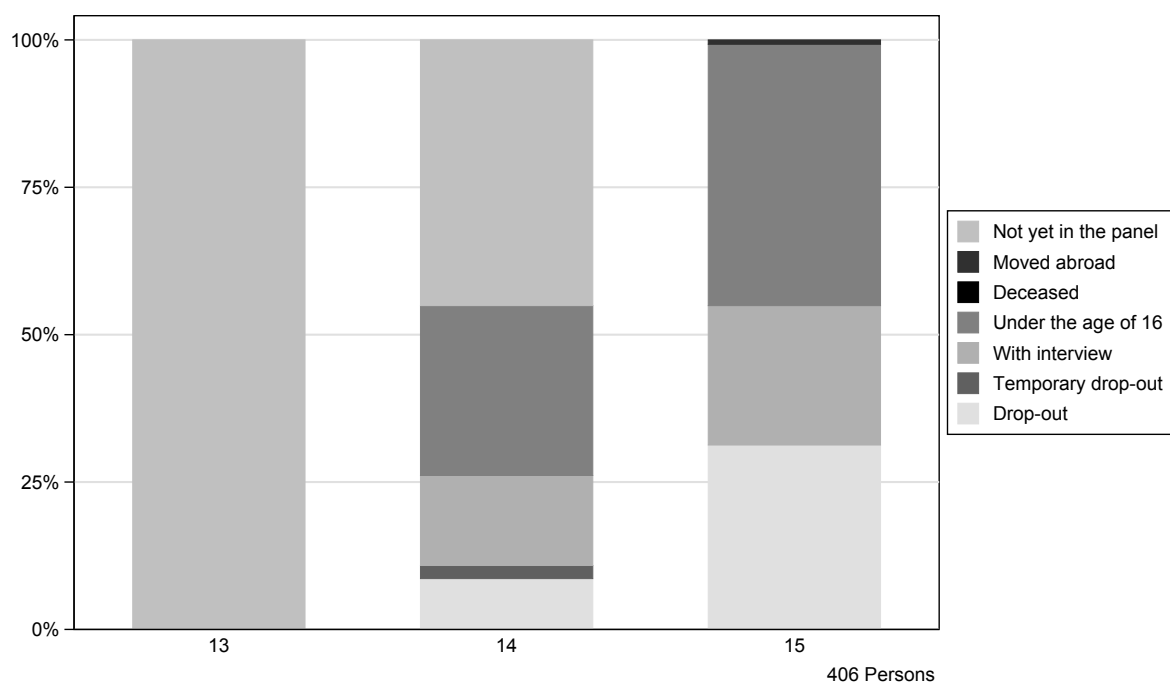


Figure 42: Entrants and their Participation Behavior (Subsample M1). Development up to Wave 3

2.4 The Risk of Survey-Related Panel Attrition

The following figures display Kaplan-Meier estimates of the survey related attrition risk (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 43, 44, 45 and 46) and some basic socio-demographic characteristics measured in the year of sampling, such as age, occupation, income, and education (Figures 47, 48, 49 and 50). 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 43), for instance, first-wave respondents from sample B have a somewhat lower probability of remaining in the survey than respondents from samples A or C. In the more recent samples D through K (Figures 44 and 45), 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.

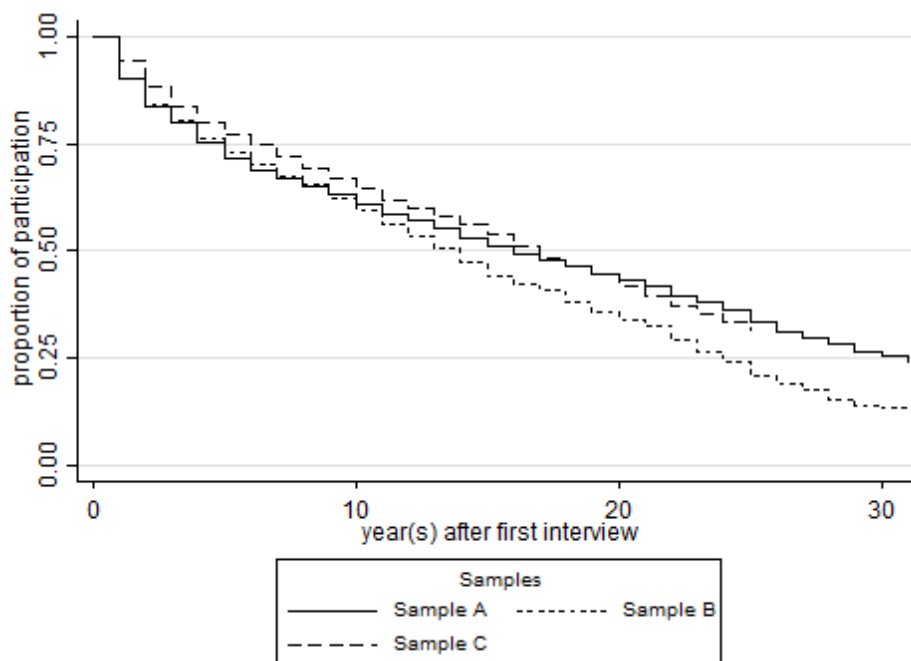


Figure 43: Successful Re-Interviewing of First-Wave Respondents by Subsamples A, B, C. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

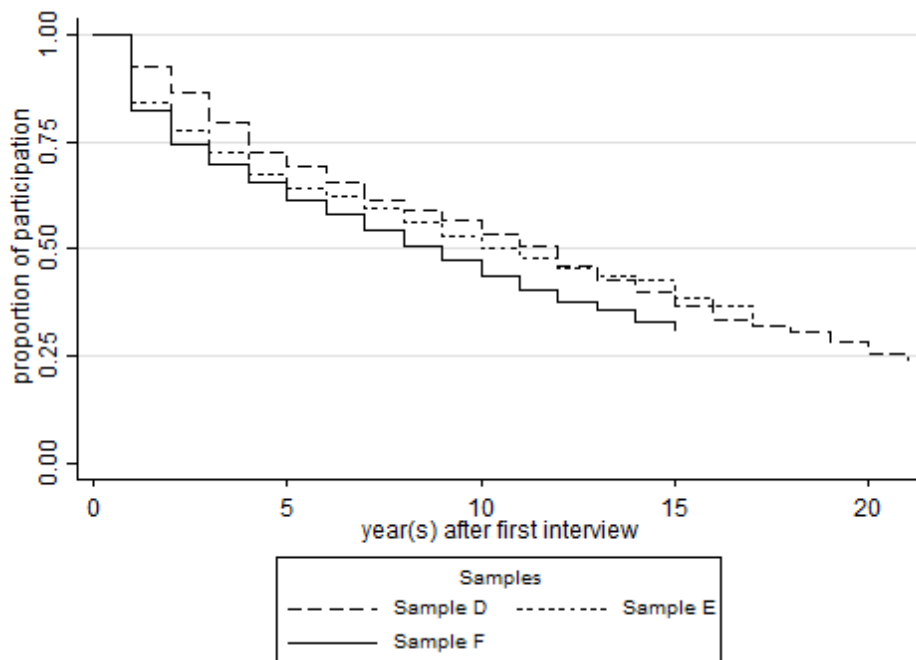


Figure 44: Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

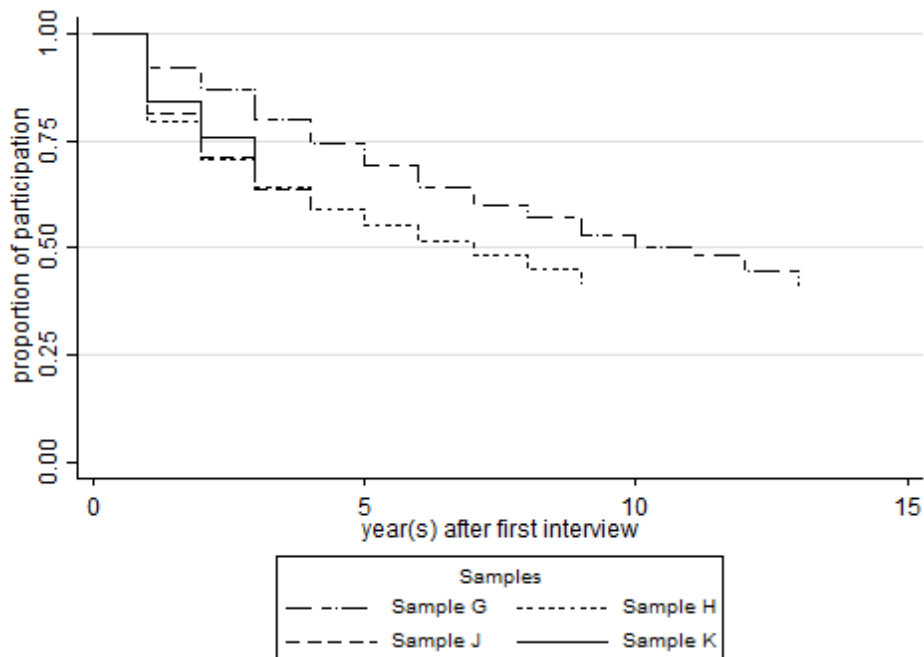


Figure 45: Successful Re-Interviewing of First-Wave Respondents by Subsamples G, H, J and K. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

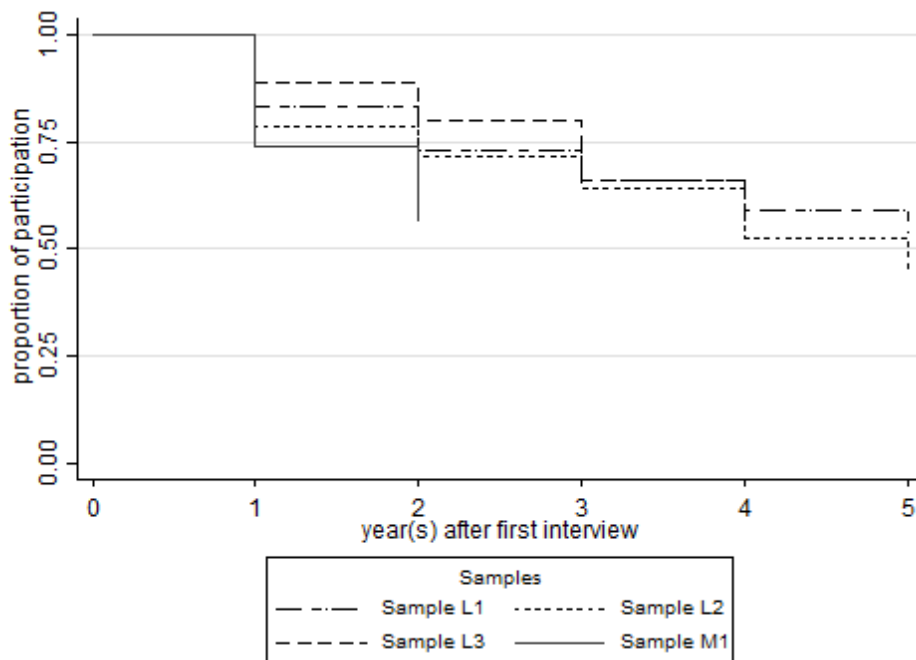


Figure 46: Successful Re-Interviewing of First-Wave Respondents by Subsamples L1, L2, L3 and M1. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

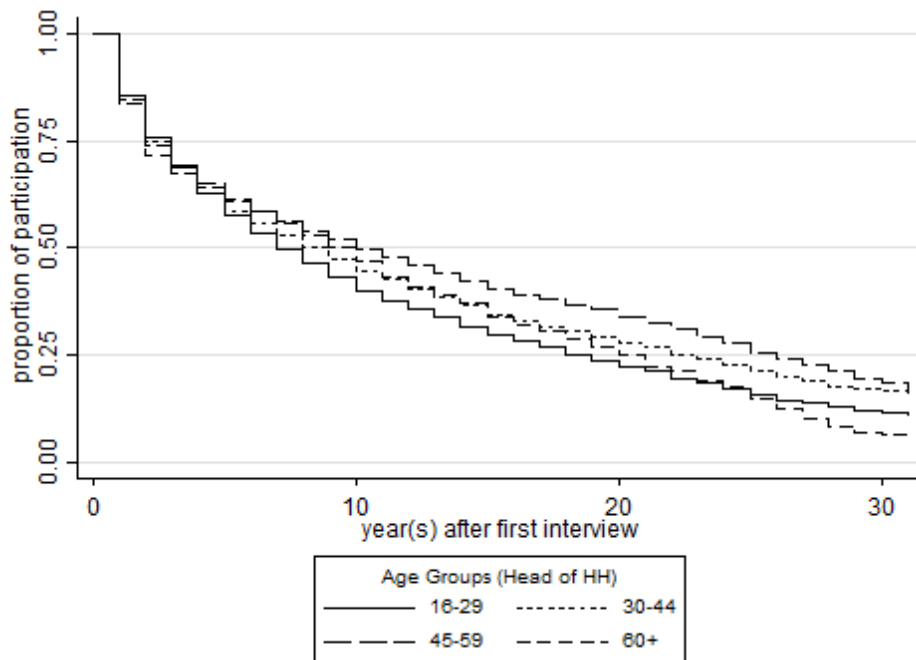


Figure 47: Successful Re-Interviewing of All First-Wave Respondents by Age Categories. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

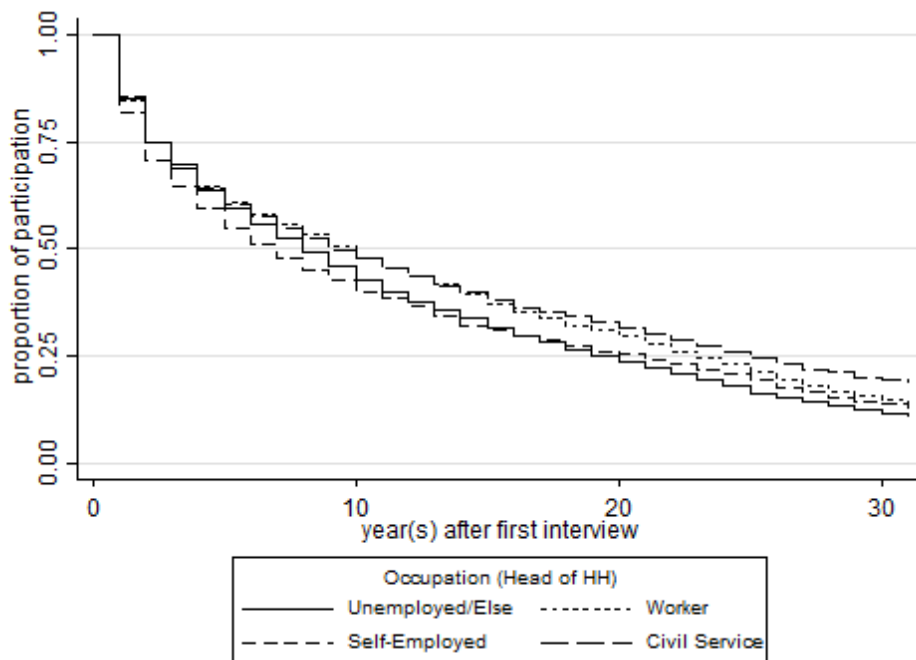


Figure 48: Successful Re-Interviewing of All First-Wave Respondents by Occupation. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

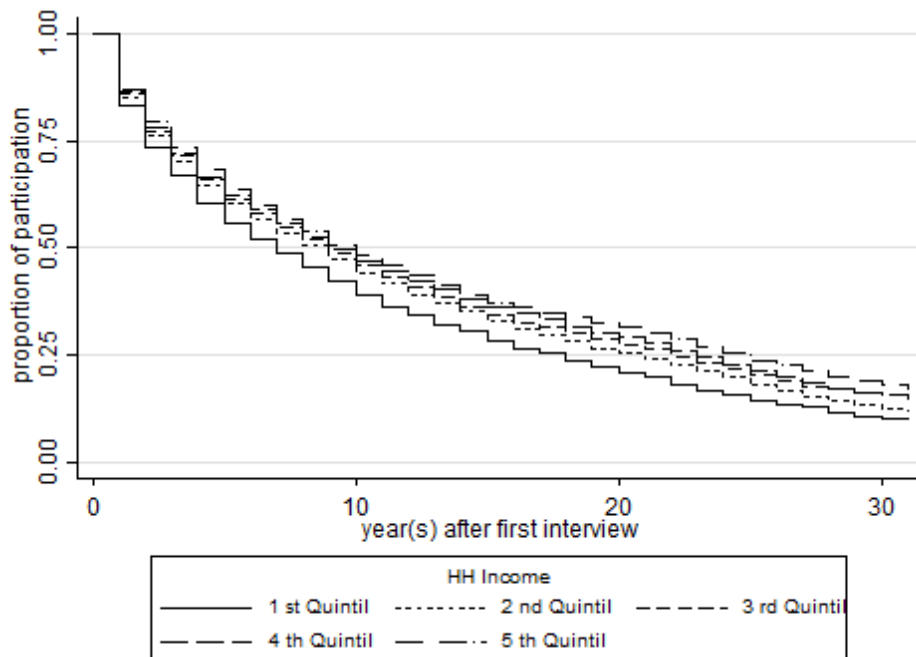


Figure 49: Successful Re-Interviewing of All First-Wave Respondents by Income Quintiles. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

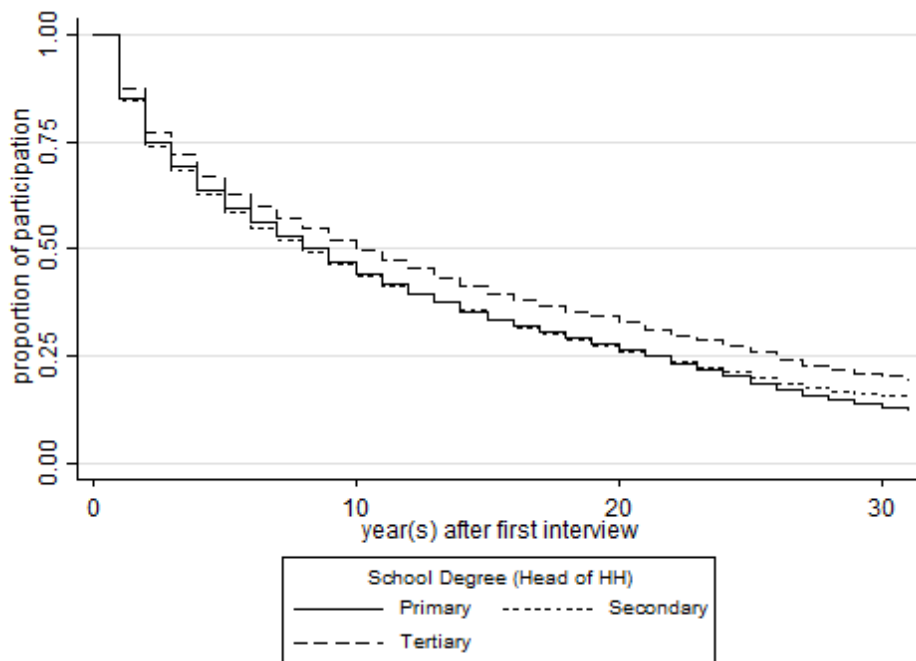


Figure 50: Successful Re-Interviewing of All First-Wave Respondents by Education. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

3 Panel Attrition Due to Unsuccessful Follow-Ups

In each panel wave, the first step in successful re-interviewing is the identification of the place of residence of households who took part in the preceding wave. The fieldwork organization of the SOEP, Kantar Public (formerly, TNS Infratest), 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 3.1 the number of households of the previous waves that need to be re-contacted and the relative frequency of successful follow-ups in sub-samples A through M1 and waves 1985 through 2015. 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.

Table 3.1: The Frequency of Households to be Re-Contacted and the Percentage of Successful Follow-Ups, Subsamples A to M1 by Year.

Year	Sample A		Sample B		Sample C		Sample D		Sample E		Sample F		Sample G		Sample H		Sample I		Sample J		Sample K		Sample L1		Sample L2		Sample L3		Sample M1	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1984	4.528	100,0	1.393	100,0																										
1985	4.681	98,3	1.370	96,4																										
1986	4.486	98,9	1.325	97,0																										
1987	4.232	99,0	1.220	98,6																										
1988	4.140	99,1	1.191	99,0																										
1989	3.984	99,0	1.157	99,0																										
1990	3.902	99,1	1.124	98,8	2.179	100,0																								
1991	3.860	99,5	1.151	99,2	2.246	98,4																								
1992	3.845	99,7	1.153	99,2	2.302	99,4																								
1993	3.867	99,2	1.172	98,5	2.227	99,0																								
1994	3.849	99,2	1.150	98,9	2.134	99,4	236	100,0																						
1995	3.784	99,5	1.108	98,9	2.110	99,5	540	100,0																						
1996	3.747	99,6	1.069	99,2	2.103	99,4	544	99,6																						
1997	3.688	99,6	1.038	99,0	2.087	99,4	541	99,2																						
1998	3.667	99,4	1.019	99,3	2.079	99,3	528	98,9	1.056	100,0																				
1999	3.631	99,6	975	99,3	2.037	99,6	498	99,3	1.089	99,4																				
2000	3.549	99,6	934	99,4	2.025	99,6	467	99,8	967	99,1	6.043	100,0																		
2001	3.463	99,5	904	99,4	2.034	99,7	454	99,0	921	99,0	6.162	98,8																		
2002	3.406	99,7	877	99,0	2.005	99,5	450	99,8	873	99,4	5.447	99,4	1.224	100,0																
2003	3.330	99,6	840	99,6	1.982	99,6	434	99,5	834	99,2	4.965	99,7	1.056	99,0																
2004	3.260	99,8	803	99,6	1.962	99,6	436	99,7	797	99,7	4.736	99,6	1.010	99,7																
2005	3.220	99,8	779	99,3	1.959	99,7	429	99,2	783	99,9	4.577	99,7	1.001	99,7																
2006	3.138	99,7	770	99,5	1.941	99,4	425	98,6	775	99,0	4.401	99,2	995	99,4	1.506	100,0														
2007	3.000	99,7	725	99,4	1.834	99,9	387	99,4	727	99,7	4.157	99,5	933	99,2	1.530	99,4														
2008	2.856	99,8	676	99,1	1.767	99,5	372	99,4	680	99,7	3.962	99,3	904	99,6	1.326	99,5														
2009	2.730	99,7	620	99,2	1.695	99,9	351	99,7	636	100,0	3.760	99,6	870	99,5	1.145	99,7	1.495	100,0												
2010	2.570	99,8	548	99,3	1.627	100,0	334	99,6	605	99,8	3.538	99,5	826	99,9	1.059	99,5	1.737	97,6			2.074	100,0	2.500	100,0						
2011	2.421	99,8	495	99,0	1.541	99,8	302	99,3	589	100,0	3.318	99,6	797	99,6	992	99,5	3.136	100,0			2.082	98,4	2.271	97,9	924	100,0				
2012	2.289	99,8	440	99,7	1.466	99,9	286	100,0	116	98,9	3.076	99,9	774	99,7	928	99,9	3.201	99,1	1.526	100,0	1.865	99,5	2.254	98,3	943	98,5				
2013	2.180	99,6	392	99,1	1.417	99,7	269	99,1	98	100,0	2.880	99,7	733	99,6	877	99,5	2.869	99,4	1.564	98,8	1.752	99,1	2.177	98,6	920	99,0	2.723	100,0		
2014	2.077	99,3	358	99,3	1.351	99,6	249	100,0	90	100,0	2.741	99,6	725	99,2	828	99,3	2.519	99,0	1.447	99,2	1.510	99,3	2.027	97,5	836	98,0	2.819	98,4		
2015	1.998	99,3	331	99,3	1.300	99,5	229	100,0	83	100,0	2.597	99,1	699	99,2	790	99,7	2.309	99,3	1.360	99,2	1.404	99,3	1.879	97,9	789	97,2	2.453	97,1		

Note: In the case of the initial wave of a sample, table entries are the number of participating households. See also Section 2.

3.2 Predicting the Probability of Successful vs. Unsuccessful Follow-Ups in the Year 2015

Based on household and interview level characteristics measured in 2014, we aim at predicting the probability of re-contacting a household relative to unsuccessful follow-up in 2015. Among a very large number of regressors that we tested in preliminary analyses, we identified a small set of variables that exert a robust effect on the probability of successful follow-ups ($p < 0.05$). Table 3.2 describes the regressors and Table 3.3 reports the subsample-specific estimates of logit models for the probability of re-contacting a household relative to unsuccessful follow-up.

Note that the estimates of regression models run for the previous waves of 1985 through 2015 are due to space restrictions not reported in the present data documentation. These can be obtained from previous attrition documentations.

Table 3.2: Definition of the Regressors of the Logit Model of Unsuccessful Follow-Ups

Variable	Label	Value
Interview Characteristics		
New HH	HH is New in SOEP	0/1
Original Sample Member	Head of HH is Original Sample Member	0/1
Head of HH New in SOEP	Head of HH New in SOEP (up to 3 Waves)	0/1
Moving In	HH Move	0/1
Phone Unknown	Telephone Number Undisclosed	0/1
CAWI-Interview	Computer Assisted Web Interviewing	0/1
High Item Nonresp.	High Item Nonresponse in Person Questionnaire of Head of HH	0/1
Temp. Drop Out Related HH	Temporary Drop Out of Related HH	0/1
Drop Out Related HH	Ultimate Drop Out of Related HH	0/1
Building, Area, and Region		
Schleswig-Holstein/Hamburg	HH Located in Schleswig-Holstein or Hamburg	0/1
H. Share Eastern EU	HH in Area with High Share of Eastern EU or Late Repatriates	0/1
High-Rise Building	HH Lives in a High-Rise Building (9 or more stories)	0/1
Financial Situation, Health		
Low Disposable Income	Disposable Income within the 1st Quartile	0/1
Sick Note	At Least one HH Member Was Certified Sick for more than 6 Weeks	0/1
Insurance Change	Head of HH Changed Health Insurance Provider in Previous Year	0/1
Work and Education		
Job Intention	Head of HH Intends to Obtain Employment in the Future	0/1
In Education	Head if HH is Currently in Education or Training	0/1
Blue-Collar Worker	Head of HH Is a Blue-Collar Worker	0/1
Same employer 2nd Q.	Duratuion Working for the Current Employer 2nd Q. (Head of HH)	0/1
Same employer 3rd Q.	Duratuion Working for the Current Employer 3rd Q. (Head of HH)	0/1
Demographic Characteristics		
Foreigner in HH	At least one Person with Non-German Nationality Lives in HH	0/1
Single HH	Single Household	0/1
Separation	At Least one HH Member Separated from His/Her Partner in the Last Year	0/1
Partner Moved In	Partner of a HH Member Moved In	0/1
In a Relationship	Head of HH is Currently in a Relationship	0/1
Death in the Family	A Death in the Family in Previous Year	0/1

Table 3.3: Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2015

	Sample A	Sample B	Sample C	Sample F	Sample G	Sample H	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1
Intercept	2.27***	1.63***	2.00***	3.50***	1.89***	1.79***	3.09***	2.74***	2.12***	1.78***	1.45***	2.11***
<u>Interview Characteristics</u>												
New HH				-1.30***	-1.10***		-1.01**			-0.83***	-0.86*	-0.89***
Original Sample Member							-0.82**					-0.49***
Head of HH New in SOEP										-0.51**		
Moving In	-0.83***			-0.97**			-1.03***	-1.50**		-0.97***		-0.98***
Phone Unknown	-0.90***			-0.73**			-0.71**	-1.03**		-0.48*	-1.44***	-0.56***
CAW-Interview										0.44*		
High Item Nonresp.							-0.62*					
Temp. Drop Out Related HH									-0.65*			
Drop Out Related HH												-0.78***
<u>Building, Area and Region</u>												
Schleswig-Holstein/Hamburg												-0.47**
H. Share Eastern EU											0.96**	
High-Rise Building							-0.51*					
<u>Financial Situation, Health</u>												
Low Disposable Income				-0.56*								
Sick Note					-0.81*							
Insurance Change									-0.71*			
<u>Work and Education</u>												
Job Intention	-0.49*											
In Education												-1.06**
Blue-Collar Worker												-0.68*
Same employer 2nd Q.				-0.79**								
Same employer 3rd Q.			-0.56**									
<u>Demographic Characteristics</u>												
Foreigner in HH							-0.70*					
Single HH				-0.70*								
Separation			-1.12***	-0.84*						-1.50***		
Partner Moved In								-1.13**				
In a Relationship										0.48*		
Death in the Family				-0.84**				-1.04*				

Note: ***p < 0.01; **p < 0.05; *p < 0.10. In Sample D and E, all households were re-contacted.

Table 3.3 – Continued from previous page

	Sample A	Sample B	Sample C	Sample F	Sample G	Sample H	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1
<i>No. of Observations</i>	1998	331	1300	2597	699	790	2309	1308	1404	1880	789	2456
<i>Log Likelihood</i>	-46.20	-12.21	-26.41	-48.48	-19.46	-13.96	-37.67	-27.25	-26.78	-85.57	-37.31	-150.87

Note: ***p <0.01; **p <0.05; *p <0.10. In Sample D and E, all households were re-contacted.

4 Panel Attrition Due to Refusals

In each panel wave, the second step in successful re-interviewing after having identified the location of 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 the death of a participant or her decision to move abroad, to generate the longitudinal weights.

4.1 The Frequency of Participation

Table 4.1 display the participation rates due to refusal by sub-sample and wave. The corresponding drop out rates can be then obtained following an analogous procedure. Note that in order to obtain this probability no distinction was made between the various types of refusals that can occur in a survey, such as unconditional refusals, refusals due to lack of time, or health problems, etc.

Table 4.1: The Frequency of Re-Contacted Households and the Percentage of Participation, Subsamples A to M by Year.

Year	Sample A		Sample B		Sample C		Sample D		Sample E		Sample F		Sample G		Sample H		Sample I		Sample J		Sample K		Sample L1		Sample L2		Sample L3		Sample M1	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1984	4.528	100,0	1.393	100,0																										
1985	4.611	89,8	1.326	89,1																										
1986	4.442	89,2	1.290	87,4																										
1987	4.194	93,2	1.204	92,7																										
1988	4.105	91,2	1.180	90,8																										
1989	3.949	92,4	1.146	91,0																										
1990	3.871	93,3	1.111	92,5	2.179	100,0																								
1991	3.842	94,0	1.143	92,4	2.213	91,7																								
1992	3.833	93,5	1.144	92,7	2.290	88,2																								
1993	3.838	93,9	1.156	92,0	2.208	89,2																								
1994	3.821	93,6	1.139	89,8	2.122	92,3	236	100,0																						
1995	3.766	93,6	1.097	89,5	2.101	92,2	540	96,7																						
1996	3.734	93,3	1.061	90,5	2.092	93,3	542	91,9																						
1997	3.674	94,1	1.029	90,5	2.076	93,5	537	89,2																						
1998	3.645	92,9	1.013	88,6	2.066	91,3	523	84,3	1.056	100,0																				
1999	3.616	92,0	969	88,5	2.030	93,3	495	85,9	1.084	81,7																				
2000	3.535	91,7	929	88,3	2.018	93,1	466	91,2	959	87,8	6.043	100,0																		
2001	3.448	91,9	899	90,0	2.028	91,2	450	88,4	913	88,8	6.100	80,5																		
2002	3.396	92,0	869	88,1	1.996	91,1	449	89,5	868	89,1	5.420	84,6	1.224	100,0																
2003	3.318	92,6	837	88,6	1.974	91,5	432	92,4	828	89,9	4.951	88,6	1.047	87,0																
2004	3.253	92,5	800	89,2	1.955	92,7	435	89,2	795	92,1	4.719	89,7	1.007	89,8																
2005	3.214	91,4	774	90,2	1.954	90,6	426	89,0	782	90,3	4.564	89,2	998	88,1																
2006	3.130	90,1	767	85,4	1.930	89,0	420	85,7	768	89,3	4.370	89,1	990	86,8	1.506	100,0														
2007	2.992	91,0	721	85,2	1.832	90,3	385	89,6	725	89,2	4.138	89,3	926	89,0	1.523	78,0														
2008	2.850	90,7	671	84,9	1.759	90,5	370	88,6	678	88,8	3.939	89,2	901	87,3	1.321	81,9														
2009	2.723	89,0	616	81,2	1.693	90,7	350	87,4	636	90,3	3.746	88,2	866	87,4	1.142	87,2	1.495	100,0												
2010	2.565	87,5	545	80,9	1.627	88,3	333	83,5	604	91,6	3.523	86,7	825	90,1	1.054	86,6	1.708	68,8	2.074	100,0	2.500	100,0								
2011	2.417	88,9	491	79,6	1.538	88,1	300	88,7	589	92,5	3.307	87,2	794	88,9	988	86,8	3.136	100,0	2.056	80,1	2.228	87,9	924	100,0						
2012	2.285	89,0	439	78,8	1.465	89,6	286	87,8	115	80,0	3.073	87,9	772	89,0	927	88,2	3.178	80,4	1.526	100,0	1.857	79,0	2.221	85,9	931	87,2				
2013	2.172	89,7	389	82,5	1.413	88,5	267	86,9	98	83,7	2.872	89,4	730	92,7	873	89,7	2.855	80,7	1.549	82,7	1.740	78,3	2.151	83,9	912	82,9	2.723	100,0		
2014	2.064	90,8	356	84,8	1.346	90,0	249	85,5	90	86,7	2.732	88,4	720	89,0	823	88,9	2.497	84,5	1.438	82,5	1.501	83,1	1.990	71,2	824	72,7	2.787	72,2		
2015	1.986	86,6	329	81,5	1.294	87,4	229	84,3	83	84,3	2.577	88,2	694	87,3	788	86,8	2.296	86,4	1.297	85,4	1.396	84,8	1.849	74,6	772	76,3	2.404	69,3		

Note: In the case of the initial wave of a sample, table entries are the number of participating households. See also Section 2.

4.2 Predicting the Probability of Re-Interviewing versus Refusal in the Year 2015

Based on the household and interview characteristics measured in the year 2014, and some regional information measured in 2015, we aim at predicting the probability of agreement vs. refusal to participate in the survey for households that were re-contacted in 2015. The individual attributes refer in many cases to the head of the household in the previous wave, but for split-off households the attributes are based on the information from 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 is aggregated at the level of households, for instance, rare events, such as the presence of individuals with an acute medical condition.

As in the case of predicting successful follow-ups, we use only model specifications where all included regressors are to be considered statistically significant (that is different from zero). The definition of the regressors is given in Table 4.2. Table 4.3 reports the subsample-specific estimates of logit models for the probability of participating relative to refusing to participate. Note again that the estimates of regression models of the previous waves 1985 through 2015 are not reported in the present documentation due to space restrictions. These can as well be found in previous attrition reports.

Table 4.2: Definition of the Regressors of the Logit Model of Refusal

Variable	Label	Value
Interview Characteristics		
New HH	HH is New in SOEP	0/1
Original Sample Member	Head of HH is Original Sample Member	0/1
Head of HH New in SOEP	Head of HH New in SOEP (up to 3 Waves)	0/1
Birth Cohort 2007	Sampled as Birth Cohort 2007	0/1
Email Known	Email Address Disclosed	0/1
Phone Unknown	Telephone Number Undisclosed	0/1
Partial Unit Nonresponse	Person(s) in HH did not Participate	0/1
High Item Nonresponse HH	High Item Nonresponse in HH Questionnaire	0/1
High Item Nonresponse P	High Item Nonresponse in Person Questionnaire of Head of HH	0/1
Many Missings Finan. Q.	No. of Item Nonresponse Above the Median of MV in Financial Questions	0/1
Temporary Drop-Out	Temporary Drop-Out of HH in Previous Year	0/1
Temp. Drop Out Related HH	Temporary Drop Out of Related HH	0/1
Drop Out Related HH	Ultimate Drop Out of Related HH	0/1
Exit Related HH	Exit of Related HH (Death/Emigration)	0/1
Interviewer Related HH	Same Interviewer in Related HH	0/1
Change of Interviewer	Change of Interviewer Between Last Waves	0/1
Mode Change	Change of Interview Mode between the Last Two Waves	0/1
CAP-Interview	Computer Assisted Personal Interview	0/1
CAW-Interview	Computer Assisted Web Interview	0/1
Long Interview	Duration of the Interview more than 20 Minutes	0/1
Short Interview	Duration of the Interview less than 10 Minutes	0/1
Late Interview	Interview Done Later than May	0/1
Youth Questionnaire	Adolescents in HH Filled out the Youth Questionnaire	0/1
Mother-Child-Questionnaire	Additional Mother-Child-Questionnaire	0/1

Table 4.2 – Continued from previous page

Variable	Label	Value
EVAMIN Control Gr.	EVAMIN-Project: Part of Control Group	0/1
EVAMIN Consent	EVAMIN-Project: Consent not Refused	0/1
EVAMIN No Consent	EVAMIN-Project: Consent Refused	0/1
Demographic Characteristics		
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
Naturalized Citizen	Head of HH Did Not Acquired German Citizenship at Birth	0/1
Foreigner in HH	At least one Person with Non-German Nationality Lives in HH	0/1
Single HH	One Person Living in HH	0/1
Family HH	4 or More Persons Live in HH	0/1
Single	Head of HH is Single	0/1
Divorced	Head of HH is Divorced	0/1
Widowed	Head of HH is Widowed	0/1
In Relationship	Head of HH Currently in a Relationship	0/1
Married	At Least one HH Member Got Married in Previous Year	0/1
Partner Outside	At Least one HH Member has Partner outside of the HH	0/1
Separation/Divorce	At Least one HH Member Separated from His/Her Partner in the Last Year	0/1
Moving In	Current Moving In HH	0/1
Death in the Family	A Death in the Family in Previous Year	0/1
Financial Situation, Real Estate and Insurance		
House Owner	Head of HH Is Owner of Dwelling	0/1
High Disposable Income	Disposable Income within the 3th Quartile	0/1
High Net HH Income	Net HH Income within the 4th Quartile	0/1
Low Net HH Income	Net HH Income within the 1st Quartile	0/1
No Investment Securities	HH Did Not Own Savings or Investment Securities Last Year	0/1
Investment Securities	HH Owned Savings or Investment Securities Last Year	0/1

Table 4.2 – Continued from previous page

Variable	Label	Value
Care Insurance	HH Member Receives Long-term Care Insurance Benefits	0/1
Work and Education		
Other Employment	Head of HH Employed, Neither Management Nor Untrained	0/1
Job Change	Head of HH Changed Job Previous Year	0/1
Same employer 1st Q.	Duration Working for the Current Employer 1st Q.	0/1
Same employer 2nd Q.	Duration Working for the Current Employer 2nd Q.	0/1
Job Intention	Head of HH Intends to Obtain Employment in the Future	0/1
Not Working	Head of HH Doesn't Work	0/1
Low Education	Head of HH Has Low Education Level	0/1
Retired	Head of HH Is Retired	0/1
Left Job	Head of HH Has Left a Job in Previous Year	0/1
Unemployed	At least one HH Member is Unemployed	0/1
Full Employment	All HH-Members Are Employed	0/1
Personality Traits, Well-Being and Other Characteristics		
Hobbies and Leisure	Head of HH Spends Much Time With Hobbies/Leisure	0/1
Energetic	Head of HH Felt Often Energetic in the Last 4 Weeks	0/1
Low Life Satisfaction	Head of HH is Dissatisfied with His/Her Life	0/1
Down and Gloomy	Head of HH Felt Often Down and Gloomy in the Last 4 Weeks	0/1
Achieved Less	Head of HH Achieved Less in the Last 4 Weeks due to Emotional Problems	0/1
Calm and Relaxed	Head of HH Felt Often Calm and Relaxed in the Last 4 Weeks	0/1
Not Worried About Health	Head of HH is Not Worried About the Own Health	0/1
Poor Health	Head of HH Has Poor Health	0/1
Party Identification	Head of HH Identifies with a Particular Political Party	0/1
Worried about Econ. Sit.	Head of HH is Worried About the Own Econ. Situation	0/1
Building, Area, and Region		
House in Bad Condition	Dwelling House in Need of Renovation	0/1
Dwelling too Small	Head of HH Thinks Apartment Is Too Small	0/1
Lack of Green Spaces	HH Affected by the Lack of Green Spaces in the Area	0/1

Table 4.2 – *Continued from previous page*

Variable	Label	Value
Storage Areas	Dwelling with Cellar/Storage Areas	0/1
Depopulation	HH Located in Area of High Depopulation	0/1
Urban Area	HH Located in Cities with More than 100,000 Inhabitants	0/1
Highrise Area	HH Located in Area with Many Multistorey Buildings	0/1
Large Apartments	HH Located in Area with Large Apartments	0/1
Small Apartments	HH Located in Area with High Share of Small Apartments	0/1
Single Family Houses	HH Located in Area with Single Family Houses	0/1
Family Households	HH Located in Area with High Share of Family Households	0/1
High Average Age	HH Located in Area with High Average Age	0/1
Low Average Age	HH Located in Area Low Average Age	0/1
Low Share Abitur-Graduates	HH Located in Area with Low Share of Abitur-Graduates	0/1
Low Voter Turnout	HH Located in Area with Low Turnout during Federal Election 2013	0/1
High Share of AFD Voters	HH Located in Area with High Share of AFD Voters	0/1
High Share of SPD Voters	HH Located in Area with High Share of SPD Voters	0/1
Low Purchasing Power	HH Located in Area with Low of Purchasing-Power (Microm)	0/1
Prosperous Region	HH Located in Prosperous Region	0/1
H. Share Islamic C.	HH Located in Area with High Share of People from Islamic Countries	0/1
Supraregional Newspapers	HH Located in Area with Affinity to Supraregional Newspapers	0/1
Baden-Wuerttemberg	HH Located in Baden-Wuerttemberg	0/1
Hessen	HH Located in Hessen	0/1
Bavaria	HH Located in Bavaria	0/1
Saxony-Anhalt	HH Located in Saxony-Anhalt	0/1
Thuringia	HH Located in Thuringia	0/1

Table 4.3: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2015

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1
Intercept	1.23***	0.02	0.74***	2.58***	0.78***	1.38***	0.54**	0.94***	1.39***	-1.40***	1.14***	-0.32*	0.31	0.59***
Interview Characteristics														
New HH	-0.44*		-0.62*							-0.81**				
Original Sample Member						-0.20**			-0.42***	-0.47**	-0.29*			-0.76***
Head of HH New in SOEP Birth Cohort 2007											-0.54*	-0.48***		
Email Known			0.19*								0.21*			
Phone Unknown	-0.50***		-0.54***			-0.55***		-0.64**	-0.54***	-0.54**		-0.57**	-1.33***	-0.65***
Part. Unit Nonresponse			-0.25*			-0.34***			-0.27**		-0.41**		-0.45***	-0.51***
High Item Nonresp. HH	-0.26***	-0.68***									-0.29*			
High Item Nonresp. Many Missings Finan. Q.											-0.28*			
Temporary Drop-Out	-1.58***					-1.40***		-2.01***	-1.73***		-1.35***		-0.75***	-0.77***
Temp. Drop Out Related HH														0.31**
Drop Out Related HH	-0.50**	-0.86*									-0.88**			
Exit of Related HH	-0.49***													
Interviewer Related HH							0.34*						0.58*	
Change of Interviewer	-0.76***	-1.01***	-0.75***	-1.69***		-0.74***	-0.89***	-0.61***	-0.53***	-0.51***	-0.65***	-0.72***	-0.47***	-0.27***
Mode Change												0.50***		
CAP-Interview							0.27*			1.78***		1.20***	0.92***	
CAW-Interview												0.97***	0.91***	
Long Interview			0.38***											
Short Interview						0.21*						-0.30*		0.22***
Late Interview						-0.21*							-0.36*	-0.20***
Youth Questionnaire														0.55***
Mother-Child-Questionnaire												0.17*		
EVAMIN Control Gr.							-0.56**			0.77***				
EVAMIN Consent	-0.22*									1.03***				
EVAMIN No Consent										0.93***				
Demographic Characteristics														
Age 25-34									-0.32***					

Note: ***p <0.01; **p <0.05; *p <0.10.

Table 4.3 – Continued from previous page

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1
Age 35-44									-0.32***	-0.54***				
Age 55-64	0.25**													
Age 65-74	0.37***					0.20*								
Naturalized Citizen						-0.33*						-0.47**		
Foreigner in HH				-2.03**		-0.32*								
Single HH			-0.29**	-1.40***		-0.39***			-0.20**		-1.12***	-0.31**		-0.39***
Family HH										0.31*				
Single											0.51***			
Divorced			0.42**											
Widowed						0.22*								
In a Relationship						-0.26*								
Married						0.43*		-0.68*						
Partner Outside							-0.40*					-0.25**		
Separation/Divorce						-0.58**								
Moving In								-0.81**						
Death in the Family								-0.48*						
Financial Situation, Real Estate and Insurance														
House Owner				-1.24**										0.17*
High Disposable Income							0.42**							0.23*
High Net HH Income						0.18*								
Low Net HH Income														0.19*
No Investment Securities	-0.26**													
Investment Securities	-0.25**								0.27*					
Care Insurance										-0.68**				
Work and Education														
Other Employment	-0.45*													
Job Change						-0.28**								-0.15*
Same employer 1st Q.		0.87***									0.40**			
Same employer 2nd Q.														-0.15*
Job Intention										-0.44**				
Not Working											0.33**			
Low Education		0.45*												
Retired		0.89**	0.31**											
Left Job						0.29*							-0.34*	
Unemployed													-0.36*	-0.22**

Note: ***p < 0.01; **p < 0.05; *p < 0.10.

Table 4.3 – Continued from previous page

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1
Full Employment							0.49***							
Personality Traits, Well-Being and Other Characteristics														
Hobbies and Leisure	-0.18*								0.15*					
Energetic	0.21**													
Low Life Satisfaction	-0.25*								-0.32**					
Down and Gloomy			0.29*											
Achieved Less						-0.22								
Calm and Relaxed							0.53***							
Not Worried About Health			0.25*											
Poor Health						-0.18*		-0.37**						
Party Identification								0.34**						0.17*
Worried about Econ. Sit.													0.40**	
Building, Area, and Region														
House in Bad Condition								0.34*			0.21*			
Dwelling too Small										0.34*	-0.28**			
Lack of Green Spaces														-0.30*
Storage Areas											-0.45**			
Depopulation								0.29*						
Urban Area		0.81**				0.16*								
Highrise Area													-0.39**	
Large Apartments											-0.29**			
Small Apartments								-0.38**	0.29***					
Single Family Houses		0.50*												
Family Households												-0.16*		
H. Average Age														0.18*
Low Average Age						0.20**								0.26***
Low Share Abitur-Graduates						-0.91*			-0.18*					
Low Voter Turnout		-0.54*												
High Share of AFD Voters		-0.64**												
High Share of SPD Voters											0.30*	-0.16*		
Low Purchasing Power											-0.25*			
Prosperous Region		0.57*												
H. Share Islamic C.									-0.28*					
Supraregional Newspapers														-0.14*
Hessen												0.40**		

Note: ***p < 0.01; **p < 0.05; *p < 0.10.

Table 4.3 – Continued from previous page

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1
Baden-Wuerttemb.										-0.35**				
Bavaria	0.22*					0.30**								-0.16*
Saxony-Anhalt														-0.44*
Thuringia														-0.82**
<i>No. of Observations</i>	1986	329	1294	133	83	2577	694	788	2296	1299	1396	1850	772	2407
<i>Log Likelihood</i>	-537.62	-115.47	-390.30	-29.05	-33.15	-700.08	-198.46	-221.80	-660.46	-372.53	-404.93	-821.35	-328.31	-1229.88

Note: ***p < 0.01; **p < 0.05; *p < 0.10.

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 the probability of staying in the SOEP in 2015 based on characteristics measured in 2014, variable `BFHBLEIB`, lends itself as a longitudinal weighting variable which itself corrects for selective attrition between waves 2014 and 2015. Tables 5.1, Table 5.2, Table 5.3 and Table 5.4 report some sub-sample specific summary statistics of the longitudinal weights in each wave.

The product of the cross-sectional weight in 2014, variable `BEHHRF`, and the longitudinal weight in 2015, variable `BFHBLEIB`, provide the raw data for the cross-sectional weight in 2015. In a final step, the post-stratification of the cross-sectional weights corrects them to meet benchmarks of known marginal distribution characteristics of the underlying population as of the year 2015. At the household level, these variables are the 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). With the integration of the migrant samples M1 in 2013 and M2 in 2015, we also consider additional information on the country of origin of respondents and year of migration. With the integration of Samples L1, L2, and L2 in 2014, we also employ more detailed information of the micro-census on household typologies. Tables 5.5 and 5.6 report sub-sample specific summary statistics of the derived cross-sectional weighting variable `BFHHRF` and in comparison all previous cross-sectional weights `AHHRF` through `BEHHRF`.

Table 5.1: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples A through D (Percentiles of \$HBLEIB up to Wave BF).

Year	Sample A				Sample B				Sample C				Sample D			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
1985	1,06	1,10	1,22	4.141	1,09	1,10	1,26	1.181								
1986	1,04	1,07	1,26	3.962	1,10	1,10	1,29	1.128								
1987	1,03	1,03	1,13	3.910	1,03	1,03	1,14	1.116								
1988	1,02	1,04	1,20	3.743	1,03	1,04	1,22	1.071								
1989	1,03	1,04	1,16	3.647	1,03	1,04	1,14	1.043								
1990	1,02	1,02	1,11	3.612	1,04	1,04	1,12	1.028								
1991	1,02	1,02	1,09	3.613	1,03	1,03	1,16	1.056	1,03	1,06	1,18	2.030				
1992	1,01	1,02	1,11	3.585	1,01	1,03	1,16	1.060	1,06	1,06	1,22	2.020				
1993	1,01	1,01	1,16	3.603	1,02	1,03	1,22	1.064	1,03	1,04	1,17	1.970				
1994	1,02	1,02	1,15	3.577	1,03	1,05	1,22	1.023	1,02	1,04	1,12	1.959				
1995	1,01	1,01	1,16	3.526	1,02	1,05	1,29	982	1,03	1,03	1,11	1.938				
1996	1,01	1,03	1,12	3.485	1,04	1,04	1,21	960	1,01	1,02	1,15	1.951	1,00	1,08	1,16	396
1997	1,01	1,02	1,13	3.458	1,02	1,04	1,29	931	1,02	1,04	1,12	1.942	1,05	1,09	1,09	340
1998	1,02	1,03	1,14	3.387	1,04	1,07	1,23	898	1,02	1,02	1,20	1.886	1,08	1,08	1,35	308
1999	1,02	1,02	1,20	3.325	1,04	1,04	1,22	858	1,01	1,03	1,10	1.894	1,05	1,05	1,27	300
2000	1,02	1,02	1,15	3.240	1,03	1,03	1,18	820	1,01	1,03	1,13	1.879	1,02	1,02	1,10	302
2001	1,02	1,02	1,18	3.168	1,02	1,02	1,23	809	1,02	1,02	1,16	1.850	1,03	1,03	1,18	286
2002	1,01	1,02	1,21	3.123	1,04	1,04	1,37	766	1,01	1,02	1,21	1.818	1,00	1,02	1,21	289
2003	1,01	1,03	1,14	3.072	1,01	1,03	1,31	742	1,01	1,03	1,14	1.807	1,01	1,01	1,09	290
2004	1,01	1,01	1,12	3.010	1,04	1,04	1,13	714	1,00	1,01	1,12	1.813	1,00	1,01	1,25	277
2005	1,02	1,02	1,16	2.937	1,05	1,05	1,17	698	1,00	1,02	1,15	1.771	1,00	1,02	1,34	273
2006	1,01	1,04	1,22	2.821	1,01	1,05	1,33	655	1,01	1,04	1,24	1.717	1,03	1,04	1,44	261
2007	1,01	1,03	1,14	2.723	1,03	1,07	1,24	614	1,00	1,03	1,15	1.654	1,01	1,04	1,12	248
2008	1,02	1,05	1,13	2.584	1,01	1,07	1,25	570	1,01	1,03	1,18	1.592	1,02	1,07	1,22	231
2009	1,02	1,05	1,25	2.423	1,01	1,05	1,60	500	1,00	1,03	1,21	1.535	1,00	1,02	1,16	217
2010	1,01	1,06	1,38	2.245	1,01	1,10	1,47	441	1,01	1,04	1,32	1.437	1,00	1,01	1,43	278
2011	1,00	1,04	1,27	2.148	1,01	1,07	1,55	391	1,01	1,05	1,24	1.355	1,01	1,02	1,28	266
2012	1,02	1,08	1,27	2.033	1,01	1,13	1,65	346	1,00	1,05	1,29	1.312	1,00	1,04	1,45	251
2013	1,01	1,06	1,25	1.949	1,01	1,09	1,58	321	1,01	1,07	1,27	1.250	1,01	1,06	1,39	232
2014	1,01	1,04	1,25	1.874	1,01	1,03	1,48	302	1,01	1,04	1,22	1.212	1,00	1,03	1,31	213
2015	1,01	1,06	1,29	1.760	1,01	1,09	1,61	268	1,02	1,07	1,37	1.131	1,00	1,02	1,63	117

Table 5.2: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples E through G (Percentiles of \$HBLEIB up to Wave BF).

Year	Sample E				Sample F				Sample G			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
1999	1,00	1,23	1,47	886								
2000	1,03	1,07	1,21	838								
2001	1,01	1,05	1,25	811	1,08	1,14	1,59	4.911				
2002	1,01	1,02	1,20	773	1,03	1,05	1,46	4.586				
2003	1,04	1,04	1,15	744	1,02	1,04	1,24	4.386	1,06	1,10	1,17	911
2004	1,00	1,01	1,08	732	1,02	1,03	1,19	4.235	1,02	1,03	1,25	904
2005	1,01	1,03	1,18	706	1,01	1,03	1,17	4.070	1,03	1,06	1,25	879
2006	1,00	1,03	1,21	686	1,01	1,03	1,29	3.895	1,00	1,04	1,31	859
2007	1,01	1,01	1,16	647	1,01	1,03	1,15	3.694	1,01	1,05	1,17	824
2008	1,00	1,01	1,19	602	1,01	1,03	1,14	3.513	1,01	1,03	1,18	787
2009	1,00	1,04	1,17	574	1,02	1,04	1,34	3.303	1,02	1,04	1,36	757
2010	1,01	1,04	1,25	553	1,01	1,05	1,40	3.055	1,00	1,01	1,23	743
2011	1,00	1,00	1,17	545	1,01	1,05	1,34	2.885	1,00	1,03	1,35	706
2012	1,05	1,24	1,66	92	1,02	1,08	1,30	2.702	1,02	1,07	1,24	687
2013	1,07	1,20	1,32	82	1,01	1,06	1,21	2.567	1,02	1,05	1,15	677
2014	1,03	1,03	1,42	78	1,02	1,05	1,25	2.414	1,01	1,07	1,32	641
2015	1,13	1,13	1,42	70	1,01	1,05	1,30	2.273	1,01	1,07	1,38	606

Table 5.3: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples H, J and K (Percentiles of \$HBLEIB up to Wave BF).

Year	Sample H				Sample J				Sample K			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
2007	1,04	1,16	1,46	1.188								
2008	1,01	1,03	1,18	1.082								
2009	1,01	1,03	1,22	996								
2010	1,01	1,04	1,37	913								
2011	1,00	1,05	1,31	858								
2012	1,00	1,03	1,36	818	1,05	1,19	1,52	2.555				
2013	1,00	1,05	1,27	783	1,03	1,13	1,36	2.305	1,04	1,15	1,47	1.281
2014	1,01	1,05	1,27	732	1,03	1,09	1,31	2.110	1,02	1,09	1,34	1.187
2015	1,01	1,09	1,26	684	1,02	1,06	1,25	1.983	1,02	1,05	1,31	1.108

Table 5.4: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples L1, L2, L3 and M1 (Percentiles of \$HBLEIB up to Wave BF).

Year	Sample L1				Sample L2				Sample L3				Sample M1			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
2011	1,10	1,20	1,46	1.647	1,03	1,12	1,37	1.958								
2012	1,04	1,16	1,58	1.467	1,03	1,11	1,35	1.907	1,01	1,10	1,37	806				
2013	1,03	1,11	1,59	1.362	1,03	1,09	1,37	1.805	1,02	1,11	1,47	750				
2014	1,03	1,12	1,45	1.247	1,09	1,25	1,69	1.416	1,10	1,24	1,80	599	1,08	1,28	1,81	2.012
2015	1,01	1,06	1,36	1.184	1,04	1,15	1,92	1.379	1,03	1,12	1,77	589	1,08	1,27	1,89	1.667

Table 5.5: Summary Statistics of the Derived Cross-Sectional Weights at the Household Level (Percentiles of \$HHRF up to Wave 32).

Year	p5	p10	p25	p50	p75	p90	p95	N
1984	431	597	3.805	4.725	5.647	7.130	8.248	5.921
1985	482	682	3.898	5.081	6.430	8.472	10.033	5.322
1986	541	752	3.598	5.301	6.838	9.281	11.116	5.090
1987	546	790	3.542	5.381	7.044	9.575	11.455	5.026
1988	537	803	3.567	5.638	7.541	10.353	12.534	4.814
1989	552	819	3.601	5.840	7.880	10.810	13.274	4.690
1990	698	1.073	2.217	4.603	7.042	9.900	12.390	6.819
1991	683	1.043	2.335	4.692	7.143	10.286	12.876	6.699
1992	670	1.028	2.339	4.660	7.135	10.529	13.653	6.665
1993	693	1.056	2.405	4.671	7.256	10.755	13.973	6.637
1994	710	1.103	2.399	4.672	7.281	11.209	14.714	6.559
1995	701	1.114	2.383	4.360	6.975	11.071	14.840	6.768
1996	740	1.162	2.386	4.366	7.019	11.381	15.319	6.699
1997	745	1.208	2.400	4.321	7.054	11.850	15.880	6.621
1998	984	1.356	2.331	3.973	6.221	9.888	13.137	7.492
1999	972	1.323	2.311	3.981	6.495	10.892	14.353	7.220
2000	803	1.100	1.758	2.524	3.568	5.084	6.521	13.082
2001	752	1.030	1.754	2.750	4.147	6.096	7.831	11.796
2002	507	657	1.222	2.553	4.192	6.513	8.262	12.320
2003	505	675	1.234	2.562	4.326	6.833	9.083	11.909
2004	492	670	1.215	2.535	4.420	7.262	9.840	11.644
2005	490	678	1.235	2.544	4.520	7.581	10.873	11.294
2006	457	650	1.267	2.393	4.140	6.883	9.734	12.361
2007	457	653	1.251	2.468	4.468	7.576	10.686	11.552
2008	458	655	1.278	2.553	4.751	8.225	11.589	10.921
2009	473	667	1.302	2.631	5.034	9.061	12.422	10.270
2010	219	362	667	1.430	3.657	7.407	11.078	13.888
2011	215	326	611	1.502	3.100	5.603	7.808	16.703
2012	218	329	635	1.630	3.161	5.705	7.594	16.397
2013	176	268	518	1.301	2.938	5.353	7.433	17.992
2014	200	311	618	1.518	3.351	6.218	8.540	15.946
2015	191	301	620	1.491	3.385	6.402	9.013	15.908

Table 5.6: Summary Statistics of the Derived Cross-Sectional Weights at the Person Level (Percentiles of \$PHRF up to Wave 32).

Year	p5	p10	p25	p50	p75	p90	p95	N
1984	397	553	1.174	4.364	5.224	6.052	6.800	16.173
1985	455	635	1.421	4.621	5.720	6.889	8.064	14.508
1986	489	678	1.532	4.682	6.025	7.589	8.999	13.804
1987	511	718	1.596	4.727	6.224	7.896	9.407	13.563
1988	489	686	1.625	4.893	6.567	8.503	10.201	12.872
1989	528	756	1.746	5.018	6.896	8.964	10.691	12.443
1990	684	1.023	1.900	3.441	6.147	8.281	10.234	18.254
1991	734	1.072	1.915	3.703	6.192	8.476	10.602	17.844
1992	781	1.139	1.998	3.742	6.301	8.723	11.108	17.429
1993	849	1.236	2.089	3.831	6.382	9.010	11.397	17.072
1994	876	1.286	2.110	3.842	6.416	9.257	12.078	16.715
1995	766	1.148	2.018	3.612	6.098	9.076	12.160	17.345
1996	800	1.193	2.025	3.641	6.143	9.422	12.767	16.944
1997	839	1.216	2.064	3.669	6.246	9.698	13.349	16.583
1998	916	1.269	2.040	3.512	5.591	8.515	11.398	18.249
1999	902	1.242	2.017	3.492	5.800	9.243	12.660	17.501
2000	724	974	1.564	2.313	3.215	4.560	5.854	30.784
2001	690	935	1.534	2.451	3.647	5.410	6.927	27.956
2002	433	596	1.056	2.193	3.717	5.825	7.632	29.101
2003	432	607	1.076	2.207	3.816	6.134	8.229	27.867
2004	428	602	1.072	2.187	3.899	6.525	8.870	26.918
2005	433	617	1.109	2.236	4.019	6.895	9.609	25.638
2006	406	585	1.108	2.175	3.681	6.305	8.845	27.442
2007	407	589	1.114	2.230	3.926	6.971	10.121	25.505
2008	418	602	1.144	2.298	4.133	7.648	11.279	23.792
2009	433	619	1.173	2.385	4.363	8.382	12.579	22.096
2010	175	276	532	1.024	2.509	5.414	8.592	35.945
2011	163	250	453	981	2.381	4.430	6.649	42.031
2012	165	249	468	1.098	2.529	4.522	6.754	40.351
2013	141	211	406	888	2.234	4.304	6.323	44.633
2014	157	243	478	1.071	2.585	4.929	7.252	38.931
2015	147	231	476	1.094	2.609	5.086	7.491	38.226

References

- Haisken-DeNew, J. and J. Frick (2001). *Desktop Companion to the Socio Economic Panel (SOEP)*.
- Kroh, M., R. Siegers, and S. Kühne (2015). *Gewichtung und Integration von Auffrischungstichproben am Beispiel des Sozio-oekonomischen Panels (SOEP)*. In: *Nonresponse Bias: Qualitätssicherung sozialwissenschaftlicher Umfragen*. Ed. by J. Schupp and C. Wolf. Wiesbaden: Springer. pp.409–444.
- Rendtel, U. (1995). *Lebenslagen im Wandel: Panelausfälle und Panelrepräsentativität*. Vol. 8. Campus Verlag.
- Schonlau, M., N. Watson, and M. Kroh (2010). “Household survey panels: how much do following rules affect sample size?” In:
- Schonlau, M., M. Kroh, N. Watson, et al. (2013). “The implementation of cross-sectional weights in household panel surveys”. In: *Statistics Surveys* 7, pp. 37–57.
- Spiess, M., M. Kroh, R. Pischner, and G. G. Wagner (2008). *On the Treatment of Non-Original Sample Members in the German Household Panel Study (SOEP) – Tracing, Weighting, and Frequencies*. SOEPpapers 98, DIW/SOEP, Berlin 2008.