

Siegers, Rainer; Belcheva, Veronika; Silbermann, Tobias

Research Report

Data documentation: Documentation of sample sizes and panel attrition in the German Socio-Economic Panel (SOEP) (1984 until 2017)

SOEP Survey Papers, No. 606

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Siegers, Rainer; Belcheva, Veronika; Silbermann, Tobias (2019) : Data documentation: Documentation of sample sizes and panel attrition in the German Socio-Economic Panel (SOEP) (1984 until 2017), SOEP Survey Papers, No. 606, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/195331>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-sa/4.0/>

SOEP Survey Papers

Series C – Data Documentation

SOEP – The German Socio-Economic Panel at DIW Berlin

2019

SOEP-Core v34 – Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2017)

Rainer Siegers, Veronika Belcheva, Tobias Silbermann

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. Stefan Liebig, DIW Berlin and Universität Bielefeld

Dr. David Richter, DIW 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:

Rainer Siegers, Veronika Belcheva, Tobias Silbermann. 2019. SOEP-Core v34 – Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2017). SOEP Survey Papers 606: Series C. Berlin: DIW/SOEP



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

© 2019 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 2017)

Rainer Siegers Veronika Belcheva Tobias Silbermann

April 4, 2019

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

Contents

Introduction	6
1 Sampling of SOEP Subsamples A to N	8
1.1 Sample A (1984)	8
1.2 Sample B (1984)	9
1.3 Sample C (1990)	10
1.4 Sample D (1994/95)	11
1.5 Sample E (1998)	12
1.6 Sample F (2000)	13
1.7 Sample G (2002)	14
1.8 Sample H (2006)	15
1.9 Sample I (2009)	16
1.10 Sample J (2011)	17
1.11 Sample K (2012)	18
1.12 Sample L1 (FiD) (2010)	19
1.13 Sample L2 (FiD) (2010)	20
1.14 Sample L3 (FiD) (2011)	21
1.15 Sample M1 (2013)	22
1.16 Sample M2 (2015)	23
1.17 Sample M3/4 (2016)	24
1.18 Sample M5 (2017)	25
1.19 Sample N (2017)	26
2 Developments in Sample Size	28
2.1 Development of the Number of Successful Interviews by Cross-Section	28
2.2 Continuance and Exit: The First Wave Gross Samples and their Participation Behavior	43
2.3 New Entrants through birth or move into SOEP Households and their Participation Behavior	46
2.4 Original Households and Split-Offs	49
2.5 The Risk of Survey-Related Panel Attrition	52
3 Panel Attrition Due to Unsuccessful Follow-Ups	57
3.1 The Frequency of Successful Follow-Ups	57
3.2 Predicting the Probability of Successful vs. Unsuccessful Follow-Ups in the Year 2017	59
4 Panel Attrition Due to Refusals	63
4.1 The Frequency of Participation	63
4.2 Predicting the Probability of Re-Interviewing vs. Refusal in the Year 2017	65
5 Margins used in the Post-Stratification Process	72
6 Summary Statistics of the Derived Longitudinal and Cross-Sectional Weights	87

List of Tables

3.1	The Frequency of Households to be Re-Contacted and the Percentage of Successful Follow-Ups, Subsamples A to N by Year.	58
3.2	Definition of the Regressors of the Logit Model of Unsuccessful Follow-Ups . .	60
3.3	Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2017	61
4.1	The Frequency of Re-Contacted Households and the Percentage of Participation, Subsamples A to N by Year.	64
4.2	Definition of the Regressors of the Logit Model of Refusal	66
4.3	Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2017	69
5.1	Marginal Distributions - Household Level	72
5.2	Margins - Household Level	78
5.3	Marginal Distributions - Person Level	80
5.4	Margins - Person Level	85
6.1	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples A through D (Percentiles of \$HBLEIB up to Wave 34).	88
6.2	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples E through G (Percentiles of \$HBLEIB up to Wave 34).	89
6.3	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples H, J and K (Percentiles of \$HBLEIB up to Wave 34).	90
6.4	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples L1, L2 and L3 (Percentiles of \$HBLEIB up to Wave 34).	90
6.5	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples M1 and M2 (Percentiles of \$HBLEIB up to Wave 34).	90
6.6	Summary Statistics of the Derived Cross-Sectional Weights at the Household Level (Percentiles of \$HHRF up to Wave 34).	91
6.7	Summary Statistics of the Derived Cross-Sectional Weights at the Person Level (Percentiles of \$PHRF up to Wave 34).	92

List of Figures

1	The Number of Successful Interviews with Persons by Subsamples A through N, Waves 1 to 34.	28
2	Comparison of Successful Interviews with Persons and Households (Subsamples A and B), Waves 1 to 34	29
3	Comparison of Successful Interviews with Persons and Households (Subsample C), Waves 1 to 28	30
4	Comparison of Successful Interviews with Persons and Households (Subsample D), Waves 1 to 23	31
5	Comparison of Successful Interviews with Persons and Households (Subsample E), Waves 1 to 20	32
6	Comparison of Successful Interviews with Persons and Households (Subsample F), Waves 1 to 18	33
7	Comparison of Successful Interviews with Persons and Households (Subsample G), Waves 1 to 16	34
8	Comparison of Successful Interviews with Persons and Households (Subsample H), Waves 1 to 12	35
9	Comparison of Successful Interviews with Persons and Households (Subsample J), Waves 1 to 7	36
10	Comparison of Successful Interviews with Persons and Households (Subsample K), Waves 1 to 6	37
11	Comparison of Successful Interviews with Persons and Households (Subsample L1), Waves 1 to 8	38
12	Comparison of Successful Interviews with Persons and Households (Subsample L2), Waves 1 to 8	39
13	Comparison of Successful Interviews with Persons and Households (Subsample L3), Waves 1 to 7	40
14	Comparison of Successful Interviews with Persons and Households (Subsample M1), Waves 1 to 5	41
15	Comparison of Successful Interviews with Persons and Households (Subsample M2), Waves 1 to 3	42
16	First-Wave Persons and their Participation Behavior. Development up to 2017	43
17	Entrants and their Participation Behavior. Development up to 2017	46
18	Proportion of First-Wave and New Households. Development up to 2017	49
19	Successful Re-Interviewing of First-Wave Respondents by Subsamples A, B, C. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	52
20	Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	53
21	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	53

22	Successful Re-Interviewing of First-Wave Respondents by Subsamples L1, L2 and L3. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	54
23	Successful Re-Interviewing of First-Wave Respondents by Subsamples M1 and M2. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	54
24	Successful Re-Interviewing of All First-Wave Respondents by Age Categories. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	55
25	Successful Re-Interviewing of All First-Wave Respondents by Occupation. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	55
26	Successful Re-Interviewing of All First-Wave Respondents by Income Quintiles. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	56
27	Successful Re-Interviewing of All First-Wave Respondents by Education. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	56

Introduction

This data documentation is meant to provide SOEP users with a general overview of the longitudinal development of the survey over the past 34 years and the derivation of weights that compensate for disproportional sampling probabilities, selective nonresponse in the first wave of each sample, as well as panel attrition.

In the first section we provide a short description of each of the SOEP samples, including structured information about the underlying target population, sampling methodology and initial fieldwork results.

In the second 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 subsamples A through K individually, the boost samples of specific family types L1-L3, the IAB-SOEP Migration Samples M1 and M2. For a general overview on the integration of enlargement and refreshment samples into the SOEP see Kroh et al. (2015a).

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).

Furthermore, the present paper gives information on the longitudinal development of the SOEP and reports descriptive figures of the participatory behavior of the original sample members, the entrance patterns of new sample members and the development of the share of original households compared to new households resulting from household splits.

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 Schonlau et al. (2013) and for a general overview, Kara et al. (2018)). 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 third 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 subsamples, age, educational, and income groups).

The fourth section reports in more detail on the occurrence of unsuccessful follow-ups to household addresses by cross-section and subsample, and subsample-specific regression models of the probability of unsuccessful follow-ups in 2017 based on the characteristics of households measured in 2016. The fifth section does the same for the second form of survey-related attrition: refusals. Documentation of panel attrition of previous panel waves can be obtained from the respective annual documentation (see, for instance, Kroh et al. (2018) for wave bg).

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 2017: `BHHBLEIB` and `BHPBLEIB`. Based on the inverse probability of observing households and persons in 2016, the staying probability in 2017, and additional post-stratification to meet benchmarks of known margins of the underlying

population in 2017, we derive the cross-sectional weights $BHHHRF$ and $BHPHRF$.

Section 6 illustrates the margins used during the post-stratification process across different waves and samples. Especially samples L1-L3 and M1-M5, that cover specific sub-populations, required a modified selection and coding of the employed margins.

The final section of this paper documents some summary statistics of the development of the longitudinal and the cross-sectional weights by subsample and wave.

1 Sampling of SOEP Subsamples A to N

1.1 Sample A (1984)

Sample A “Residents in the Federal Republic of Germany” is one of the two initial samples of the SOEP and covers private households with a household head, who does not belong to one of the main foreigner groups of “guestworkers” (i.e. Turkish, Greek, Yugoslavian, Spanish or Italian households).

Key Facts

Sampling Design	Multistage stratified sampling procedure based on the ADM-Design ¹		
<i>first stage</i>	Stratification: federal states (NUTS 1) governmental regions (NUTS 2) municipality size		
	Clustering: 585 primary sampling units (PSUs)		
<i>second stage</i>	Random walk in each PSU Selected unit: household		
Sample Size ²	households	persons (thereof children)	
	NET	4,524	11,366 (2,290)
	GROSS	7,430	
Field Period	February to October 1984		
Initial Survey Mode	Paper-and-Pencil Interviewing (PAPI), possibility for self-completion		
Number of Interviewers	592		
Initial Response Rate ³	60.9%		
Initial Weighting Factor	Average	SD	min / max
	5,491	1,402	3,358 / 11,041
Further Readings			

Infratest Sozialforschung (2011). *SOEP 1984 – Methodenbericht zum Befragungsjahr 1984 (Welle 1) des Sozio-oekonomischen Panels*. SOEP Survey Papers 1, DIW/SOEP, Berlin 2011.

¹ADM is the “Arbeitsgemeinschaft Deutscher Marktforschungsinstitute” (Working Group of the German Market Research Institutes). For more information, see <https://www.adm-ev.de/persnlich-muendlichebefragungen/>

²The net sample includes households and persons with complete or partial interview. The gross sample comprises also the non-participating households, excluding those that were classified as “quality neutral non-response” (e.g. invalid addresses, deaths, moving abroad).

³AAPOR Response Rate Definition RR2, see AAPOR (2016).

1.2 Sample B (1984)

Sample B “Foreigners in the Federal Republic of Germany” is one of the two initial Samples of the SOEP and covers private households with a Turkish, Greek, Yugoslavian, Spanish or Italian household head. Compared to Sample A the population of Sample B is oversampled.

Key Facts

Sampling Design	Multistage stratified sampling procedure using the registers of foreigners in each county (Ausländerregister der Landkreise)		
<i>first stage</i>	Stratification: federal states (NUTS 1) governmental regions (NUTS 2) number of foreigners of the respective nationality		
	Clustering: 241 PSUs (random selection of PSUs independent for each nationality)		
<i>second stage</i>	Random selection of addresses in each PSU		
	Selected unit: person		
Sample Size	households	persons (thereof children)	
	NET	1,393	4,807 (1,638)
	GROSS	2,045	
Field Period	April to October 1984		
Initial Survey Mode	Paper-and-Pencil Interviewing (PAPI)		
Number of Interviewers	253		
Initial Response Rate	68.1%		
Initial Weighting Factor	Average	SD	min / max
	820	574	89 / 4,347
Further Readings	<p>Infratest Sozialforschung (2011). <i>SOEP 1984 – Methodenbericht zum Befragungsjahr 1984 (Welle 1) des Sozio-oekonomischen Panels</i>. SOEP Survey Papers 1, DIW/SOEP, Berlin 2011.</p>		

1.3 Sample C (1990)

Sample C “German Residents in the German Democratic Republic (GDR)” covers persons in private households in which the household head was a citizen of the GDR.

Key Facts

Sampling Design	Multistage stratified sampling procedure based on GDR-Master-Sample designed by Infratest in cooperation with the Department for Social Research of the Radio of GDR ⁴		
<i>first stage</i>	Stratification: counties (NUTS 3) municipality size		
	Clustering: 330 PSUs		
<i>second stage</i>	Random walk in each PSU with start addresses drawn from the central residents' data base		
	Selected unit: household		
Sample Size	households	persons (thereof children)	
	NET	2,179	6,044 (1,591)
	GROSS	3,404	
Field Period	May to July 1990		
Initial Survey Mode	Paper-and-Pencil Interviewing (PAPI), possibility for self-completion		
Number of Interviewers	215		
Initial Response Rate	64.0%		
Initial Weighting Factor	Average	SD	min / max
	3,103	1,734	367 / 19,102
Further Readings	<p>Infratest Sozialforschung (2011). <i>SOEP 1990/91 – Methodenbericht Ostdeutschland zu den Befragungsjahren 1990-1991 (Welle 1/2 – Ost) des Sozio-ökonomischen Panels</i>. SOEP Survey Papers 14, DIW/SOEP, Berlin 2011.</p>		

⁴In German: Abteilung Soziologische Forschung des Rundfunks der DDR.

1.4 Sample D (1994/95)

Sample D “Immigrants” covers private households in which at least one household member had moved from abroad to West Germany after 1984. It mainly consists of ethnic Germans migrating from Eastern Europe to Germany. This sample includes two subsamples that were drawn independently in 1994 (D1) and in 1995 (D2).

The fieldwork organization sampled a small number of households of Sample D ($N=98$) drawing on a respondent-driven sampling procedure. In these 98 cases, inclusion probabilities cannot be derived directly and we thus do not assign weights to these households.

Key Facts

Sampling Design

first stage The migrant households were identified in representative surveys of the German population in 1992 (D1) and 1994 (D2).

second stage D1: Additional respondent-driven sampling units
D2: Random selection of households for two of the migrant groups (ethnic German immigrants from GDR and from Eastern Europe)

Selected unit: household

Sample Size

	households	persons (thereof children)
	D1 D2	D1 D2
NET	236 295 ⁵	719 (248) 905 (283)
GROSS	307 385	

Field Period

January to March 1994 (D1) and January to April 1995 (D2)

Initial Survey Mode

Paper-and-Pencil Interviewing (PAPI), possibility for self-completion

Number of Interviewers

83 (1994) 206 (1995)

Initial Response Rate

76.9% (D1) 76.6% (D2)

Initial Weighting Factor

Average SD min / max

(in 1995) 3,906 1,717 1,699 / 9,855

Further Readings

Infratest Sozialforschung (2011). *SOEP 1994 – Methodenbericht Zuwanderer-Befragung (Teilstichprobe D1) zum Befragungsjahr 1994 (Welle 11) des Sozio-oekonomischen Panels*. SOEP Survey Papers 26, DIW/SOEP, Berlin 2011.

Infratest Sozialforschung (2011). *SOEP 1995 – Methodenbericht Zuwanderer-Befragung II (Zweitbefragung D1, Erstbefragung D2) zum Befragungsjahr 1995 (Welle 12) des Sozio-oekonomischen Panels*. SOEP Survey Papers 28, DIW/SOEP, Berlin 2011.

Rendtel, U., M. Pannenberg and S. Daschke (1997). *Die Gewichtung der Zuwanderer-Stichprobe des Sozio-oekonomischen Panels (SOEP)*. In: Vierteljahrshefte zur Wirtschaftsforschung, Duncker & Humblot, Berlin, Vol. 66. Iss. 2, pp. 271-286.

⁵213 cases in Sample D do not meet the requirements of the SOEP sampling design. These cases are interviewed, but do not receive valid weights.

1.5 Sample E (1998)

Sample E “Refreshment I” is the first sample that was designed to be representative for all private households in both East and West Germany. It is the first of several regular refreshment samples drawn to increase the overall size of the SOEP, compensate for panel-attrition and cover population changes, e.g. due to migration.

It is also the first sample in which the Computer-Assisted Personal Interviewing (CAPI) was implemented. Interviews in Samples A-D at this time were completely conducted using Paper-and-Pencil Interviewing (PAPI). To study mode effects, households of sample E were randomly allocated to CAPI and PAPI mode.

With the data distribution of 2012, parts of sample E have been extracted into the SOEP Innovation Sample.

Key Facts

Sampling Design	Multistage stratified sampling procedure based on the ADM-Design		
<i>first stage</i>	Stratification: federal states (NUTS 1) governmental regions (NUTS 2) municipality size		
	Clustering: 125 PSUs		
<i>second stage</i>	Random walk in each PSU Selected unit: household		
Sample Size	households	persons (thereof children)	
	NET	1,056	2,376 (466)
	GROSS	1,969	
Field Period	April to September 1998		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI) and Paper-and-Pencil Interviewing (PAPI)		
Number of Interviewers	130		
Initial Response Rate	53.6%		
Initial Weighting Factor	Average	SD	min / max
	35,568	18,294	14,827 / 205,099
Further Readings	<p>Infratest Sozialforschung (2011). <i>SOEP 1998 – Methodenbericht Erstbefragung der Stichprobe E zum Befragungsjahr 1998 (Welle 15) des Sozio-oekonomischen Panels</i>. SOEP Survey Papers 33, DIW/SOEP, Berlin 2011.</p> <p>Projektgruppe Das Sozio-oekonomische Panel (DIW) (1998). <i>Funktion und Design einer Ergänzungstichprobe für das Sozio-oekonomische Panel (SOEP)</i>. DIW Discussion Papers 163, Berlin 1998.</p> <p>Schräpler, J.-P., J. Schupp and G. G. Wagner (2006). <i>Changing From PAPI to CAPI – A longitudinal Study of Mode Effects Based on an Experimental Design</i>. DIW Discussion Papers 593, Berlin 2006.</p>		

1.6 Sample F (2000)

Sample F “Refreshment II” covers private households in Germany and substantially increases the sample size of the SOEP. Experience with the previous samples has shown that migrant households display lower response probabilities. This is why households with at least one adult not having the German nationality were oversampled in Sample F.

Key Facts

Sampling Design	Multistage stratified sampling procedure based on the ADM-Design		
<i>first stage</i>	Stratification: federal states (NUTS 1) governmental regions (NUTS 2) counties (NUTS 3) municipality size		
	Clustering: 985 PSUs		
<i>second stage</i>	Random walk in each PSU Oversampling of “non-German” households Selected unit: household		
Sample Size	households	persons (thereof children)	
	NET	6,043	13,871 (2,991)
	GROSS	11,862	
Field Period	March to October 2000		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI) and Paper-and-Pencil Interviewing (PAPI), possibility for self-completion		
Number of Interviewers	671		
Initial Response Rate	50,9%		
Initial Weighting Factor	Average	SD	min / max
	6,364	2,224	2,376 / 18,861
Further Readings	<p>Infratest Sozialforschung (2011). SOEP 2000 – Methodenbericht erste Welle der SOEP Stichprobe F zum Befragungsjahr 2000 (Welle 17) des Sozio-ökonomischen Panels. SOEP Survey Papers 37, DIW/SOEP, Berlin 2011.</p>		

1.7 Sample G (2002)

The 2002 Sample G “High Income” covers private households in Germany with a monthly income of at least DM 7,500 (EUR 3,835), which - due to the lack of an adequate sampling frame - were identified using a telephone screening procedure. From Wave 2 in 2003 onwards, only households with a net monthly income of at least EUR 4,500 were interviewed further.

Key Facts

Sampling Design	The households were selected from a representative telephone survey of the German population in 2001 that consists of nearly 99,000 telephone interviews.		
<i>first stage</i>	Selection of the households with corresponding income from the master sample that agreed to participate in a following study		
<i>second stage</i>	Stratification according income and region (east/west)		
Sample Size	households	persons (thereof children)	
	NET	1,224	3,364 (693)
	GROSS	2,493	
Field Period	March to July 2002		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI) and Paper-and-Pencil Interviewing (PAPI), possibility for self-completion		
Number of Interviewers	276		
Initial Response Rate	49,1%		
Initial Weighting Factor	Average	SD	min / max
	2,084	953	983 / 9,757
Further Readings			

Infratest Sozialforschung (2011). *SOEP 2002 – Methodenbericht Sondererhebung Hocheinkommensstichprobe zum Befragungsjahr 2002 (Welle 19) des Sozio-oekonomischen Panels*. SOEP Survey Papers 44, DIW/SOEP, Berlin 2011.

Infratest Sozialforschung (2011). *SOEP 2003 – Methodenbericht zweite Welle der Sondererhebung Hocheinkommensstichprobe zum Befragungsjahr 2003 (Welle 20) des Sozio-oekonomischen Panels*. SOEP Survey Papers 47, DIW/SOEP, Berlin 2011.

1.8 Sample H (2006)

Sample H “Refreshment III” covers private households in Germany. For the first time in a SOEP subsample, all households were interviewed in the computer-assisted personal interview mode (CAPI).

Key Facts

Sampling Design	Multistage stratified sampling procedure based on the ADM-Design		
<i>first stage</i>	Stratification: federal states (NUTS 1) governmental regions (NUTS 2) municipality size		
	Clustering: 250 PSUs		
<i>second stage</i>	Random walk in each PSU		
	Selected unit: household		
Sample Size	households	persons (thereof children)	
	NET	1,506	3,239 (623)
	GROSS	3,747	
Field Period	March to July 2006		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	243		
Initial Response Rate	40,2%		
Initial Weighting Factor	Average	SD	min / max
	26,443	13,453	9,024 / 12,8852
Further Readings	<p>TNS Infratest Sozialforschung (2011). <i>SOEP 2006 – Methodenbericht Erstbefragung der Ergänzungsstichprobe H zum Befragungsjahr 2006 (Welle 23) des Sozio-oekonomischen Panels</i>. SOEP Survey Papers 57, DIW/SOEP, Berlin 2011.</p>		

1.9 Sample I (2009)

Sample I “Innovation Sample” covers private households in Germany. A disproportional sampling design was implemented in order to increase the number of migrant households in the SOEP. In order to do so, an analysis of family names –“onomastic procedure” – was applied. In 2012, Sample I was completely transferred to SOEP-IS, which is why it is excluded in terms of weighting. The cases are nevertheless integrated in SOEP waves Z and BA (2009 and 2010), however, without valid weighting factors.

Key Facts

Sampling Design	Multistage stratified sampling procedure based on the ADM-Design	
<i>first stage</i>	Stratification: federal states (NUTS 1) governmental regions (NUTS 2) municipality size	
	Clustering: 250 PSUs	
<i>second stage</i>	Random walk for address listing in each PSU	
	Oversampling of migrant households such that the share of migrants for each PSU is doubled	
	Selected unit: household	
Sample Size	households	persons (thereof children)
	NET	1,495 3,052 (620)
	GROSS	4,743
Field Period	September 2009 to January 2010	
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)	
Number of Interviewers	233	
Initial Response Rate	31.5%	
Further Readings		

TNS Infratest Sozialforschung (2012). *SOEP 2009 – Methodenbericht Innovationssample zum Befragungsjahr 2009 (Welle 26) des Sozio-oekonomischen Panels (Erstbefragung Stichprobe I)*. SOEP Survey Papers 73, DIW/SOEP, Berlin 2012.

Schröder, M., D. Saßenroth, J. Körtner, M. Kroh, and J. Schupp (2013). *Experimental Evidence of the Effect of Monetary Incentives on Cross-Sectional and Longitudinal Response: Experiences from the Socio-Economic Panel (SOEP)*. SOEPpapers 603, DIW/SOEP, Berlin 2013.

Pforr, K., M. Blohm, A. G. Blom, B. Erdel, B. Felderer, M. Fräßdorf, K. Hajek, S. Helmschrott, C. Kleinert, A. Koch, U. Krieger, M. Kroh, S. Martin, D. Saßenroth, C. Schmiedeberg, E.-M. Trüdinger, and B. Rammstedt (2015). “Are Incentive Effects on Response Rates and Nonresponse Bias in Large-scale, Face-to-face Surveys Generalizable to Germany? Evidence from Ten Experiments”. In: *Public Opinion Quarterly* 79.3, 740–768.

1.10 Sample J (2011)

Sample J “Refreshment IV” covers private households in Germany. Again, a disproportional sampling design was implemented in order to increase the number of migrant households in the SOEP.

Key Facts

Sampling Design	Multistage stratified sampling procedure based on the ADM-Design		
<i>first stage</i>	Stratification: federal states (NUTS 1) governmental regions (NUTS 2) municipality size		
	Clustering: 307 PSUs		
<i>second stage</i>	Random walk for address listing in each PSU Oversampling of migrant households ⁶ such that the share of migrants for each PSU is doubled Selected unit: household		
Sample Size		households	persons (thereof children)
	NET	3,136	6,308 (1,147)
	GROSS	9,492	
Field Period	March to October 2011		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	338		
Initial Response Rate	33.0%		
Initial Weighting Factor	Average	SD	min / max
	12,592	6,166	1,950 / 49,307

Further Readings

TNS Infratest Sozialforschung (2012). *SOEP 2011 – Methodenbericht zum Befragungsjahr 2011 (Welle 28) des Sozio-oekonomischen Panels*. SOEP Survey Papers 108, DIW/SOEP, Berlin 2012.

Kroh, M., K. Käppner and S. Kühne (2014). *Sampling, Nonresponse, and Weighting in the 2011 and 2012 Refreshment Samples J and K of the Socio-Economic Panel*. SOEP Survey Papers 260, DIW/SOEP, Berlin 2014.

⁶Identification of potentially migrant households using onomastic procedure.

1.11 Sample K (2012)

Sample K “Refreshment V” covers private households in Germany.

Key Facts

Sampling Design	Multistage stratified sampling procedure based on the ADM-Design		
<i>first stage</i>	Stratification: federal states (NUTS 1) governmental regions (NUTS 2) municipality size		
	Clustering: 126 PSUs		
<i>second stage</i>	Random walk for address listing in each PSU		
	Selected unit: household		
Sample Size	households	persons (thereof children)	
	NET	1,526	3,036 (563)
	GROSS	4,397	
Field Period	March to October 2012		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	304		
Initial Response Rate	34.7%		
Initial Weighting Factor	Average	SD	min / max
	26,053	10,305	5,343 / 80,336

Further Readings

TNS Infratest Sozialforschung (2013). *SOEP 2012 - Methodenbericht zum Befragungsjahr 2012 (Welle 29) des Sozio-oekonomischen Panels*. SOEP Survey Papers 144, DIW/SOEP, Berlin 2013.

Kroh, M., K. Käppner and S. Kühne (2014). *Sampling, Nonresponse, and Weighting in the 2011 and 2012 Refreshment Samples J and K of the Socio-Economic Panel*. SOEP Survey Papers 260, DIW/SOEP, Berlin 2014.

1.12 Sample L1 (FiD) (2010)

Sample L1 “Cohort Sample”⁷ covers private households in Germany, in which at least one household member is a child that was born between January 2007 and March 2010. Again, migrants identified by an “onomastic procedure” are oversampled.

Key Facts

Sampling Design	Multistage stratified sampling procedure based on information from local registration offices (Einwohnermeldeämter)		
<i>first stage</i>	Stratification: federal states (NUTS 1) governmental regions (NUTS 2) municipality size		
	Clustering: 159 PSUs		
<i>second stage</i>	Random selection of children in the respective cohort in each PSU provided by the local registration offices, stratified by municipality size		
	Oversampling of migrant households ⁸ such that the share of migrants for each PSU is doubled		
	Selected unit: child in the respective cohort		
Sample Size	households	persons (thereof children)	
	NET	2,074	7,670 (3,900)
	GROSS	5,286	
Field Period	June to October 2010		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	204		
Initial Response Rate	39.2%		
Initial Weighting Factor	Average	SD	min / max
	935	576	75 / 3,494
Further Readings	<p>TNS Infratest Sozialforschung (2010). <i>“Familien in Deutschland” (FiD) 2010 Methodenbericht: Anlage und Ergebnisse der FiD-Stichproben.</i> München 2011.</p> <p>Schröder, M., R. Siegers, K. Spieß (2013). <i>“Familien in Deutschland” - FiD.</i> Schmollers Jahrbuch: Vol. 133, No. 4, pp. 595-606.</p>		

⁷Sample L1 (as well as L2 and L3) was part of the SOEP-related study “Familien in Deutschland” (FiD), which was later integrated into the SOEP in 2014. As part of an evaluation project of the Federal Ministry for Family Affairs, Senior Citizens, Women and Youth (BMFSFJ) and the Federal Ministry of Finance (BMF) the study focused on public benefits in Germany for married people and families. Therefore, the survey instruments of waves BA to BD differ in some parts from those of the other samples.

⁸Identification of potentially migrant addresses using onomastic procedure and information on the citizenship.

1.13 Sample L2 (FiD) (2010)

Sample L2 “Family Types I” covers private households in Germany that meet at least one of the following criteria regarding their household composition: single parents, low income families and large families with three or more children. Similar to Sample G we face the problem that the eligible sub-population is relatively small and an adequate sampling frame is lacking. So again, a preceding telephone screening procedure identifies eligible households.

Key Facts

Sampling Design	Persons in potentially eligible households were identified in representative face-to-face and telephone surveys of the German population following the ADM-Design. Telephone screening (CATI-Screening) was then conducted in order to verify the eligibility and willingness of the households to participate. Selected unit: person		
Sample Size	households	persons (thereof children)	
	NET	2,500 ⁹	8,838 (4,611)
	GROSS	3,281	
Field Period	March to June 2010		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	343		
Initial Response Rate	76.2%		
Initial Weighting Factor	Average	SD	min / max
	1,596	1,035	213 / 7,702
Further Readings			

TNS Infratest Sozialforschung (2010). *“Familien in Deutschland” (FiD) 2010 Methodenbericht: Anlage und Ergebnisse der FiD-Stichproben.* München 2011.

Schröder, M., R. Siegers, K. Spieß (2013). *“Familien in Deutschland” - FiD.* Schmollers Jahrbuch: Vol. 133, No. 4, pp. 595-606.

⁹During the fieldwork in wave 1,237 households were identified not to be part of the target population and thus do not receive valid weights.

1.14 Sample L3 (FiD) (2011)

Sample L3 “Family Types II” covers private households in Germany that meet at least one of the following criteria regarding their household composition: single parents or large families with three or more children. It is conducted analogical to Sample L2 in order to increase the number of cases in these sub-populations.

Key Facts

Sampling Design	Persons in potentially eligible households were identified in representative face-to-face and telephone surveys of the German population following the ADM-Design. Telephone screening (CATI-Screening) was then conducted to verify the eligibility and willingness of the households to participate. Selected unit: person		
Sample Size	households	persons (thereof children)	
	NET	924 ¹⁰	3,579 (2,092)
	GROSS	1,144	
Field Period	March to June 2011		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	250		
Initial Response Rate	80.8%		
Initial Weighting Factor	Average	SD	min / max
	2,359	1,582	468 / 12,146
Further Readings			

[TNS Infratest Sozialforschung \(2011\). “Familien in Deutschland” \(FiD\) 2011 Methodenbericht: Anlage und Ergebnisse der FiD-Stichproben. München 2011.](#)

[Schröder, M., R. Siegers, K. Spieß \(2013\). “Familien in Deutschland” - FiD. Schmollers Jahrbuch: Vol. 133, No. 4, pp. 595-606.](#)

¹⁰During the fieldwork of the first wave, 9 households were identified not to be part of the target population and thus do not receive valid weights.

1.15 Sample M1 (2013)

The 2013 “IAB-SOEP Migration Sample” (M1) was jointly planned and conducted by the Institute for Employment Research (IAB) in Nuremberg and the German Socio-Economic Panel (SOEP) at DIW Berlin. Register data of the Federal Employment Agency (FEA), the so-called Integrated Employment Biographies (IEB), were used as a sampling frame. The target population consists of individuals in the register as of 31.12.2011 who a) immigrated to Germany since 1995 as well as b) second-generation migrants born after 1976 in Germany.

Key Facts

Sampling Design	Multistage stratified sampling design based on the IEB database		
<i>first stage</i>	Stratification: federal states (NUTS 1) county type (urban/rural)		
	Clustering: 250 PSUs proportional to number of migrants ¹¹ in each stratum		
<i>second stage</i>	Simulated random walk algorithm in each PSU		
	Disproportional address sampling according to country of origin and migration generation		
	Selected unit: person		
Sample Size	households	persons (thereof children)	
	NET	2,723	7,445 (2,481)
	GROSS	11,051	
Field Period	May to November 2013		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	232		
Initial Response Rate	35,0% ¹²		
Initial Weighting Factor	Average	SD	min / max
	1,564	1,540	63 / 9,117
Further Readings			

TNS Infratest Sozialforschung (2014). *Methodenbericht zum IAB-SOEP-Migrationssample 2013*. SOEP Survey Papers 217, DIW/SOEP, Berlin 2014.

Kroh, M., S. Kühne, J. Goebel and F. Preu (2015). *The 2013 IAB-SOEP Migration Sample (M1): Sampling Design and Weighting Adjustment*. SOEP Survey Papers 271, DIW/SOEP, Berlin 2015.

Eisnecker, P. S., K. Erhardt, M. Kroh, and P. Trübswetter (2017). *The Request for Record Linkage in the IAB-SOEP Migration Sample*. SOEP Survey Papers 291, DIW/SOEP, Berlin 2017.

Eisnecker, P. S. and M. Kroh (2017). “The Informed Consent to Record Linkage in Panel Studies: Optimal Starting Wave, Consent Refusals, and Subsequent Panel Attrition”. In: *Public Opinion Quarterly* 81.1, 131-143

¹¹Identification of target persons using information on nationality, FEA measures and onomastic procedure.

¹²Including the 1,145 households that were screened out and not taken into further consideration.

1.16 Sample M2 (2015)

The 2015 “IAB-SOEP Migration Sample” (M2) aimed for the collection of information on households with recent migrants, that is, individuals who immigrated to Germany between 2009 and 2013. Similar to the M1 sample, register data of the Federal Employment Agency was used as a sampling frame.

Key Facts

Sampling Design	Multistage stratified sampling design based on the IEB database		
<i>first stage</i>	Stratification: federal states (NUTS 1) county type (urban/rural) proportion of migrants in each PSU		
	Clustering: 125 PSUs proportional to the number of target population members ¹³ in each stratum		
<i>second stage</i>	Disproportional address sampling in each PSU according to country of origin		
	Selected unit: person		
Sample Size	households	persons (thereof children)	
	NET	1,096	2,638 (927)
	GROSS	6,008	
Field Period	May to December 2015		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	143		
Initial Response Rate	32,6% ¹⁴		
Initial Weighting Factor	Average	SD	min / max
	901	741	54 / 3,760
Further Readings	<p>Kühne, S. and M. Kroh (2017). <i>The 2015 IAB-SOEP Migration Study M2: Sampling Design, Nonresponse, and Weighting Adjustment</i>. SOEP Survey Papers 473, DIW/SOEP, Berlin 2017.</p>		

¹³Identified by the year they entered the IEB and former and current citizenship.

¹⁴Including the 863 households that were screened out and not taken into further consideration.

1.17 Sample M3/4 (2016)

The 2016 “IAB-BAMF-SOEP Refugee Survey” (Samples M3 and M4) is a joint project of the Institute for Employment Research (IAB), the Research Centre of the Federal Office for Migration and Refugees (BAMF-FZ) as well as the Socio-economic Panel (SOEP). The target population of the samples consists of households with individuals who arrived in Germany between January 2013 and January 2016 and applied for asylum or were hosted as part of specific programs of the federal states (irrespective of their asylum procedure and their current legal status).

The first part of the sample (M3) was financed with funds from the research budget of the Federal Employment Agency (BA) allocated to the IAB. Sample M4 was funded by the Federal Ministry of Education and Research (BMBF) and has a focus on refugee families.

Key Facts

Sampling Design	Multistage stratified sampling design based on the German Central Register of Foreigners (AZR)		
<i>first stage</i>	Stratification: federal states (NUTS 1) county type (urban/rural) Clustering: 99 PSUs (M3) / 95 PSUs (M4)		
<i>second stage</i>	Disproportional address sampling in each PSU according to country of origin, current legal status, age and gender Selected unit: person		
Sample Size	households	persons (thereof children)	
	NET	3,320	9,965 (5,438)
	GROSS	6,848	
Field Period	June to December 2016		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	164		
Initial Response Rate	48,5%		
Initial Weighting Factor	Average	SD	min / max
	142	192	4 / 2,967

Further Readings

Kroh, M., H. Brücker, S. Kühne, E. Liebau, J. Schupp, M. Siebert, and P. Trübswetter (2016). *Das Studiendesign der IAB-BAMF-SOEP-Befragung von Geflüchteten*. SOEP Survey Papers 365, DIW/SOEP, Berlin 2016.

Kroh, M., S. Kühne, J. Jacobsen, M. Siebert, and R. Siegers (2017). *Sampling, Nonresponse, and Integrated Weighting of the 2016 IAB-BAMF-SOEP Survey of Refugees (M3/M4) – revised version*. SOEP Survey Papers 477, DIW/SOEP, Berlin 2017.

1.18 Sample M5 (2017)

Sample M5 is both an enlargement and a refreshment of the former sub-samples M3 and M4 which are known as the IAB-BAMF-SOEP Survey of Refugees. Whereas the target population of M3 and M4 are all people that immigrated to Germany between January 2013 and January 2016 and appeared in the Central Register of Foreigners up to April 2016, M5 adds two new aspects: First, people that immigrated to Germany between January 2013 and January 2016 and made a claim for asylum after April 2016 until January 2017 (refreshment) and, second, people who immigrated to Germany between February 2016 and December 2016 and making a claim for asylum until January 2017 (enlargement). The sampling is similar to sampling of M3 and M4 and we propose, for substantial analyses, to use all three sub-samples jointly. By using all sub-samples together they are representative for people immigrating to Germany and applied for asylum or people who were hosted as part of specific programs of the federal states (irrespective of their asylum procedure and their current legal status).

Key Facts

Sampling Design	Multistage stratified sampling design based on the German Central Register of Foreigners (AZR)		
	<i>first stage</i>	Stratification: federal states (NUTS 1) county type (urban/rural)	
		Clustering: 99 PSUs	
	<i>second stage</i>	Disproportional address sampling in each PSU according to country of origin, current legal status, gender, and target population (refreshment vs. enlargement)	
		Selected unit: person	
Sample Size		households	persons (thereof children)
	NET	1,519	4,161 (1,909)
	GROSS	2,915	
Field Period	June to October 2017		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	30		
Initial Response Rate	52.1%		
Initial Weighting Factor	Average	SD	min / max
	138	162	5/2189
Further Readings	<p>Jacobsen, J., M. Kroh, S. Kühne, J. A. Scheible, R. Siegers, and M. Siegert (2019). <i>Supplementary of the IAB-BAMF-SOEP Survey of Refugees in Germany (M5) 2017</i>. SOEP Survey Papers 605, DIW/SOEP, Berlin 2019.</p>		

1.19 Sample N (2017)

Cases of Sample N were initially drawn in the context of the international *Project in Assessment of Adult Skills and Competencies* (PIAAC) in 2012 that was initiated by the OECD¹⁵. The survey of the German subsample was carried out by GESIS and the target population of PIAAC 2012 Germany consisted of adults from 16 through 65 years that lived in Germany (on the reference date of 1 December 2011). The fieldwork in 2012 resulted in a net sample of 5,319 persons. Participants were then transferred into the PIAAC-L panel study¹⁶, which followed the concept of "Anchor Persons", meaning that only original PIAAC sample members were followed in subsequent waves. The waves of PIAAC-L surveyed not only the PIAAC anchor persons, but other household members as well and already introduced items similar to those of the SOEP. The respective waves were conducted in the years 2014 (3,758 anchor interviews), 2015 (3,263) and 2016 (2,967), of which 2,811 anchor persons have agreed to be transferred into the SOEP. Finally, Sample N is based on respondents that took part in the last wave of PIAAC-L in 2016 and gave consent to be transferred into the SOEP.

Key Facts

Sampling Design ¹⁷	Two-staged stratified and clustered sampling procedure based on information from local registration offices (Einwohnermeldeämter)		
<i>first stage</i>	Stratification: federal states administrative regions districts county type (rural/urban)		
<i>second stage</i>	Clustering: 277 PSUs systematic random sampling in each PSU Selected unit: person ¹⁸		
Sample Size ¹⁹	households	persons (thereof children)	
	NET	2,378 ²⁰	4,807 (1,037)
	GROSS	3,447	
Field Period	March to August 2017		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	287		
Initial Response Rate	67.1%		
Initial Weighting Factor	Average	SD	min / max
	14,016	11,455	1,755 / 157,875

¹⁵A detailed description of the international PIAAC survey can be found in OECD (2016).

¹⁶For more detailed information on the respective waves please see the corresponding Technical Reports listed under *Further Readings*.

¹⁷The sampling design outlined here refers to the initial sample of PIAAC Germany in 2012.

¹⁸The households of the initially drawn persons for PIAAC 2012 provided the basis for PIAAC-L and Sample N, by also interviewing other household members, after giving their consent to participate.

¹⁹The numbers in this paragraph refer to the actual Sample N of the SOEP. For information concerning the respective PIAAC and PIAAC-L samples see the literature listed below.

²⁰64 of these households will be realised the first time in wave 2 of Sample N.

Further Readings

Zabal, A., S. Martin, N. Massing, D. Ackermann, S. Helmschrott, I. Barkow, and B. Rammstedt (2014). *PIAAC Germany 2012. Technical report*. Münster: Waxmann.

OECD, 2nd Edition (2016). *Technical Report of the Survey of Adult Skills (PIAAC)*. Not yet published.

Zabal, A., S. Martin, and B. Rammstedt (2016). *PIAAC-L data collection 2014: technical report; follow-up to PIAAC Germany 2012*. GESIS Papers, 2016|17. Köln: GESIS - Leibniz-Institut für Sozialwissenschaften.

Zabal, A., S. Martin, and B. Rammstedt (2017). *PIAAC-L data collection 2015: technical report*. GESIS Papers 2017|29, Köln: GESIS - Leibniz-Institut für Sozialwissenschaften.

Martin, S., A. Zabal, and B. Rammstedt (2018). *PIAAC-L data collection 2016: technical report*. GESIS Papers 2018|05, Köln: GESIS - Leibniz-Institut für Sozialwissenschaften.

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 the entrance of new sample members by birth / moving into SOEP households and their participation behavior, (2.4) reporting share of original households in relation to new households from splits and (2.5) 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 database.

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

The following figures display the number of successful interviewed cases at the household and individual level.

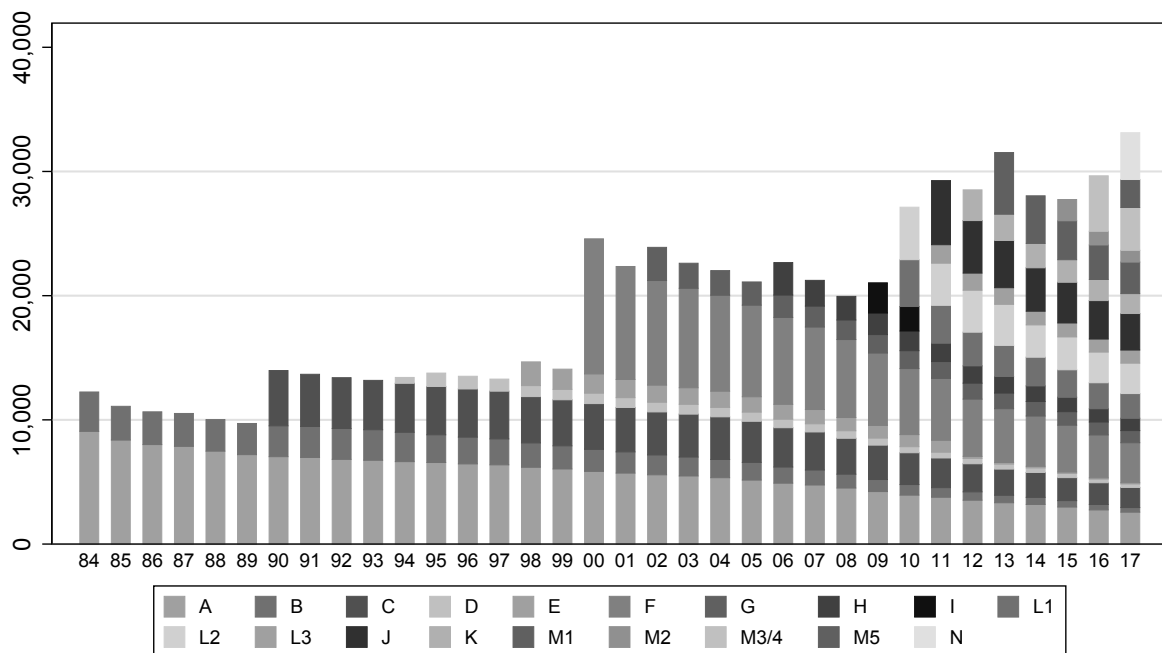


Figure 1: The Number of Successful Interviews with Persons by Subsamples A through N, Waves 1 to 34.

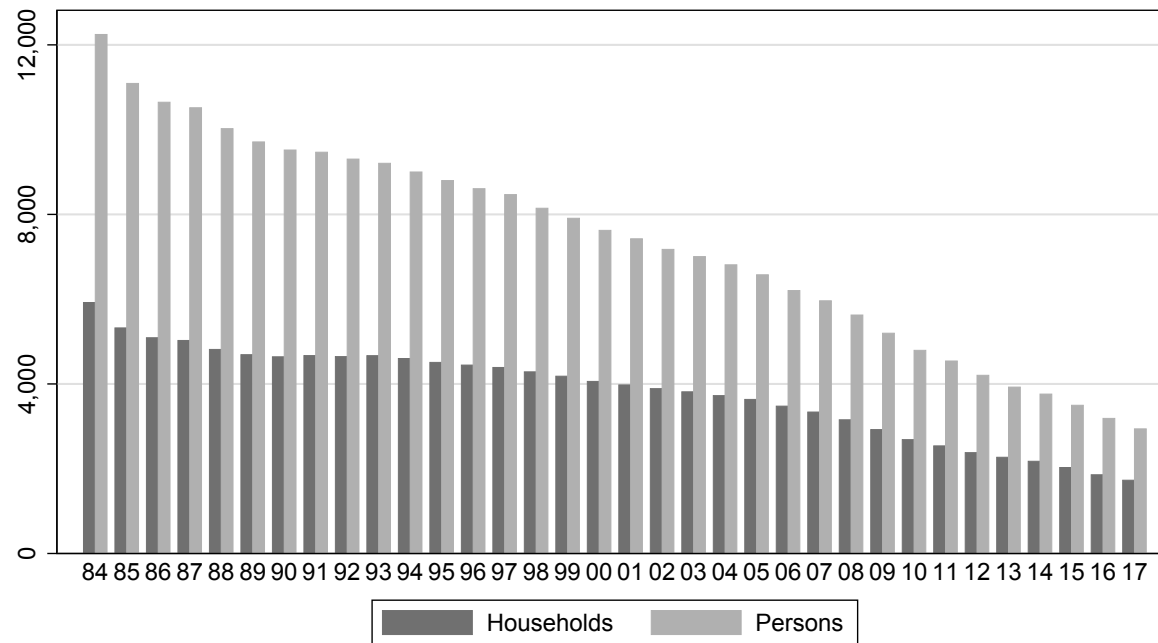


Figure 2: Comparison of Successful Interviews with Persons and Households (Subsamples A and B), Waves 1 to 34

Year	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
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
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

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Persons	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	3,187	2,940
Households	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	1,857	1,729

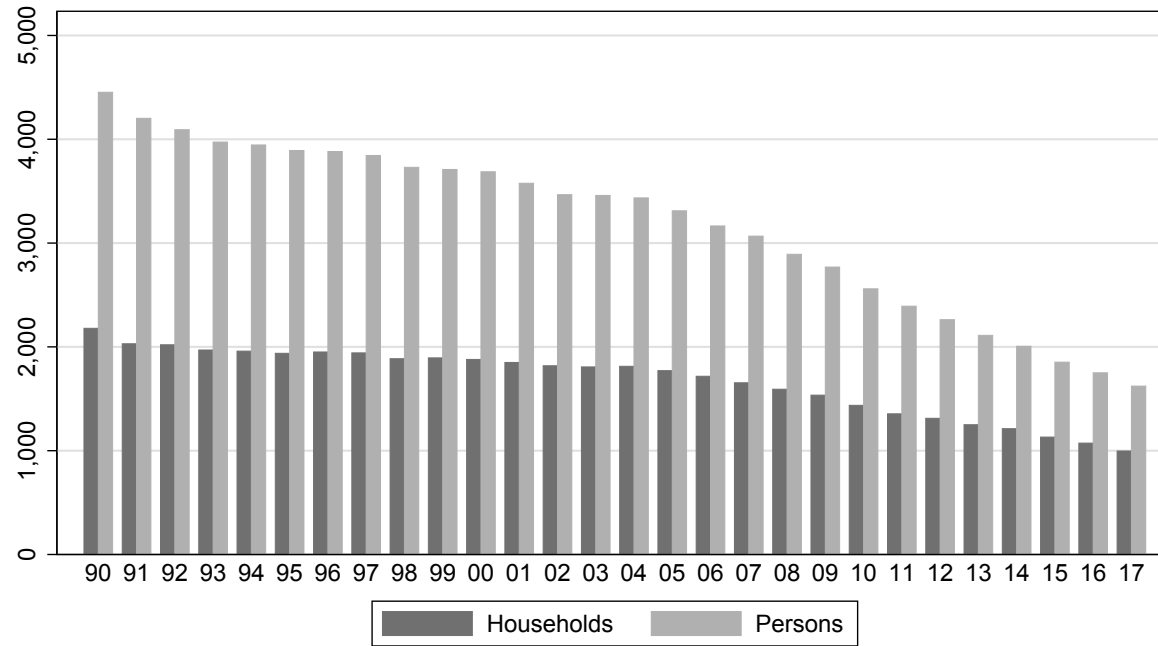


Figure 3: Comparison of Successful Interviews with Persons and Households (Subsample C), Waves 1 to 28

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
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	3,459
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

Year	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Persons	3,435	3,311	3,165	3,067	2,892	2,769	2,559	2,392	2,262	2,111	2,006	1,853	1,750	1,622
Households	1,813	1,771	1,717	1,654	1,592	1,535	1,437	1,355	1,312	1,250	1,212	1,131	1,073	997

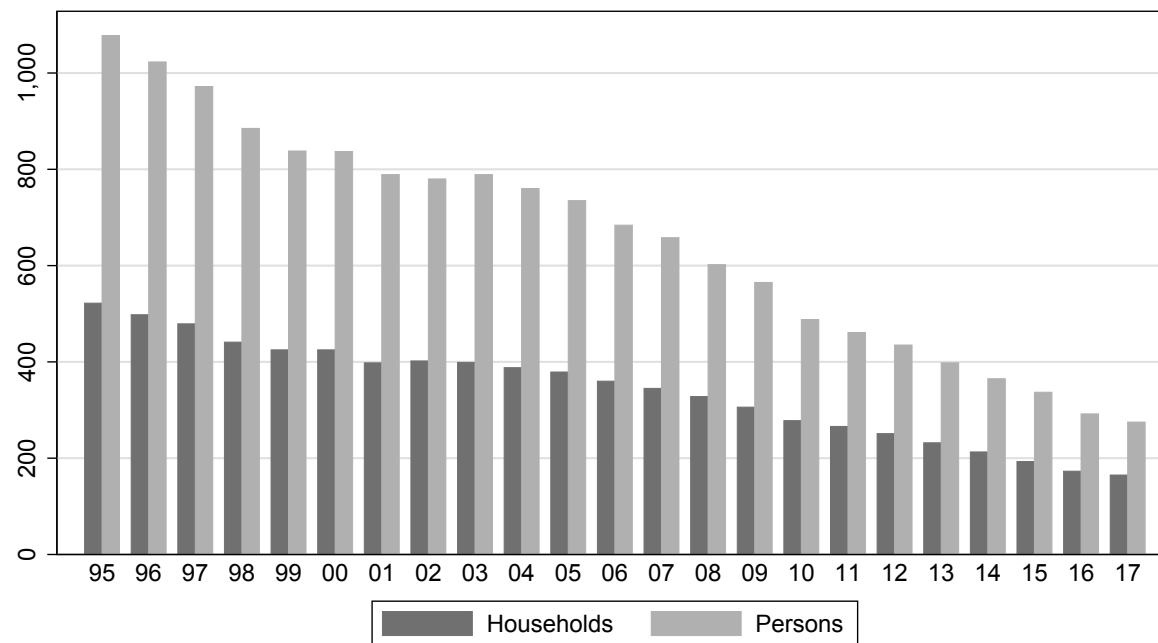


Figure 4: Comparison of Successful Interviews with Persons and Households (Subsample D), Waves 1 to 23

Year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Persons	1,078	1,023	972	885	838	837	789	780	789	760	735	684
Households	522	498	479	441	425	425	398	402	399	388	379	360

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Persons	658	602	565	488	461	435	398	365	337	292	275
Households	345	328	306	278	266	251	232	213	193	173	165

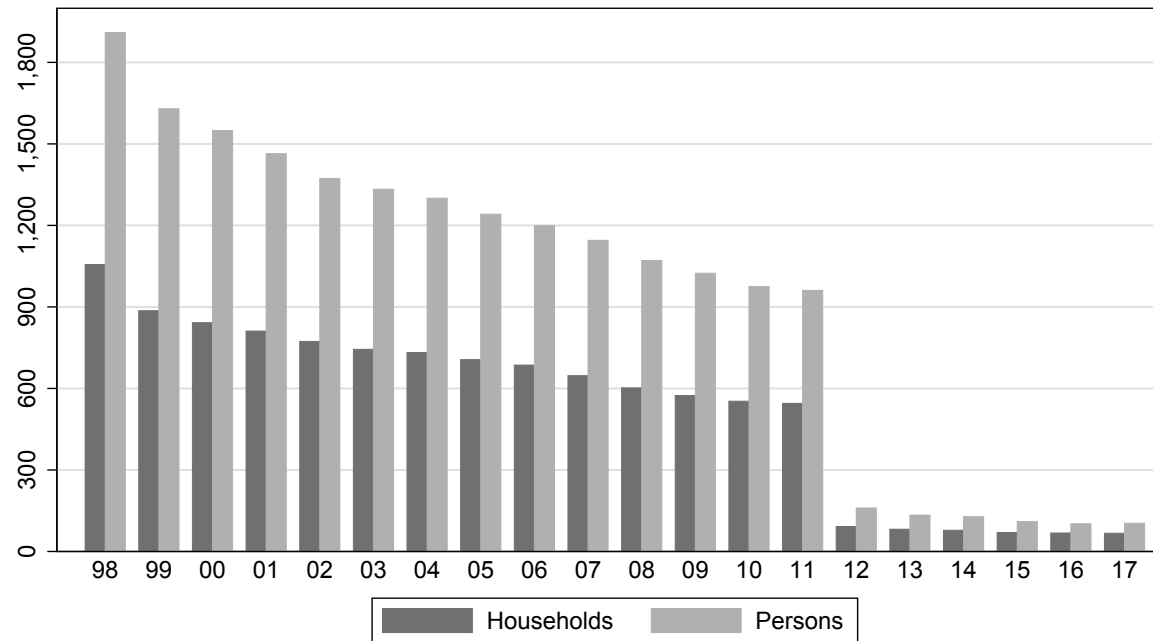


Figure 5: Comparison of Successful Interviews with Persons and Households (Subsample E), Waves 1 to 20²¹

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Persons	1,910	1,629	1,549	1,464	1,373	1,333	1,300	1,241	1,199	1,145
Households	1,056	886	842	811	773	744	732	706	686	647

Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Persons	1,071	1,024	975	961	160	134	128	110	102	104
Households	602	574	553	545	92	82	78	70	68	67

²¹In 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.

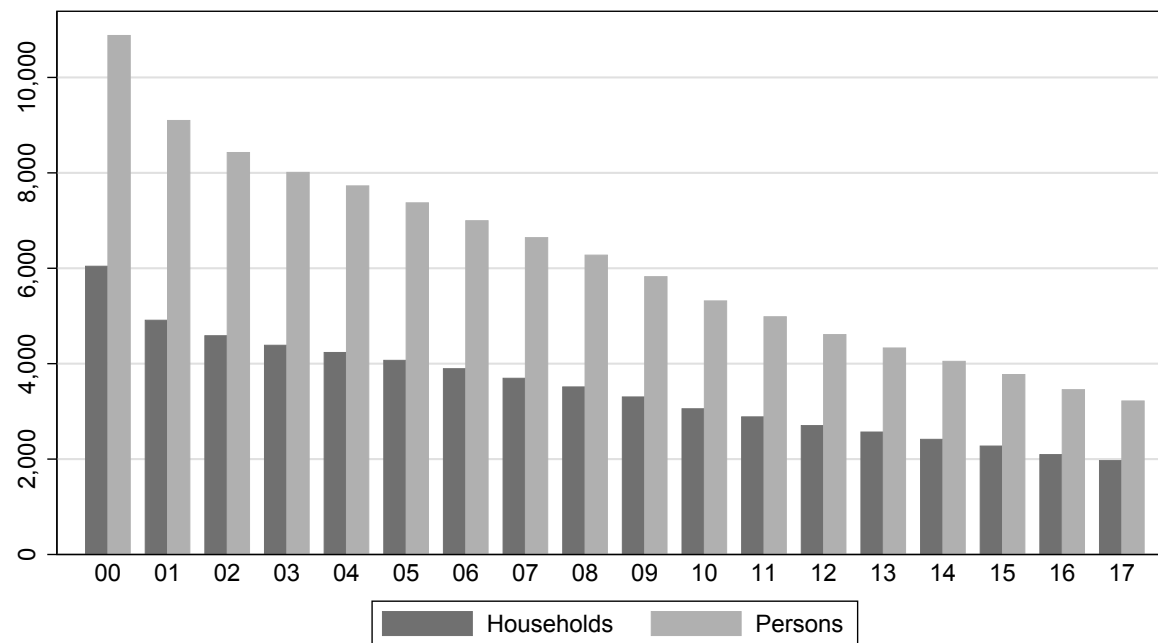


Figure 6: Comparison of Successful Interviews with Persons and Households (Subsample F), Waves 1 to 18

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
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	3,455	3,219
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	2,094	1,968

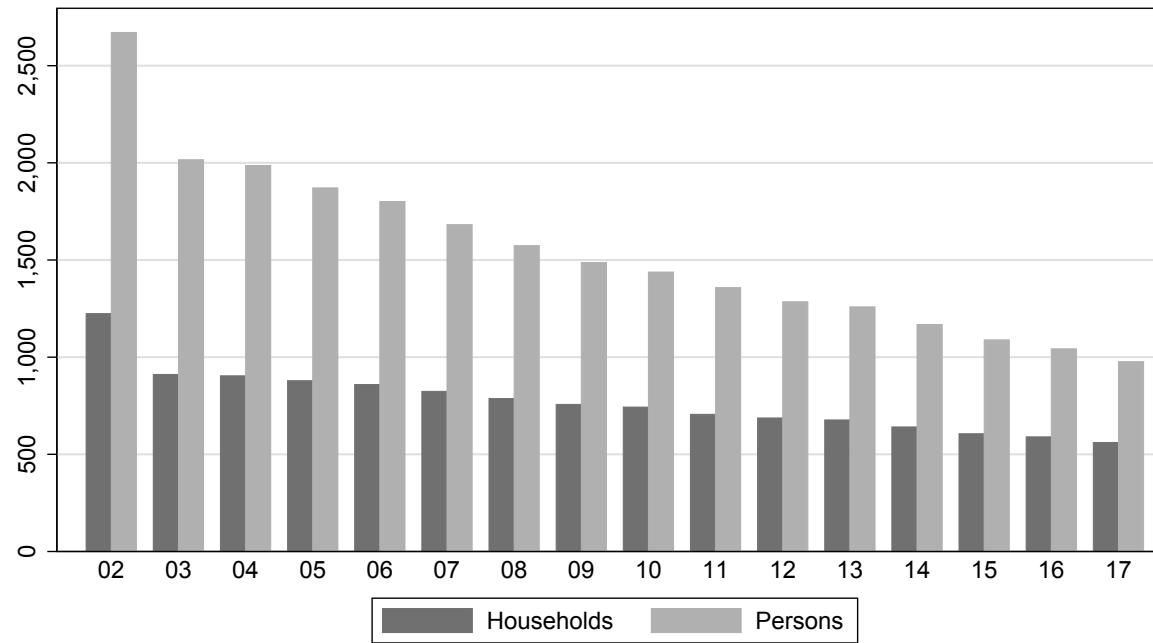


Figure 7: Comparison of Successful Interviews with Persons and Households (Subsample G), Waves 1 to 16²²

Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
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	1,043	977
Households	1,224	911	904	879	859	824	787	757	743	706	687	677	641	606	590	561

²²In the second wave the target population was changed: a higher income threshold resulted in a smaller number of observations in 2003.

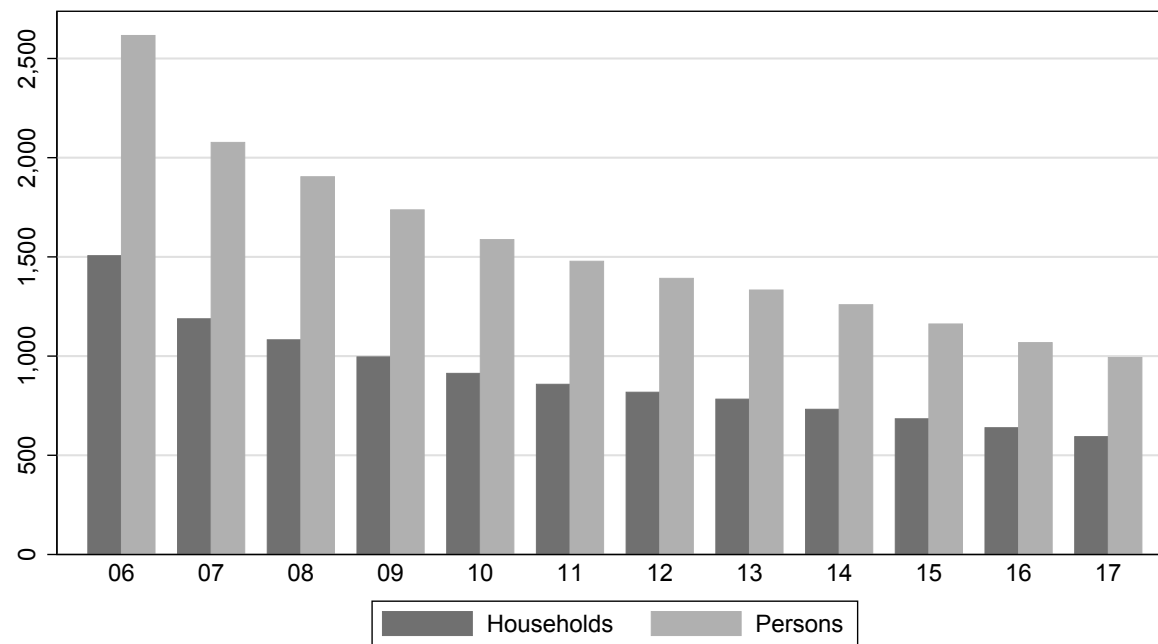


Figure 8: Comparison of Successful Interviews with Persons and Households (Subsample H), Waves 1 to 12

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Persons	2,616	2,077	1,904	1,737	1,587	1,478	1,392	1,333	1,259	1,162	1,068	993
Households	1,506	1,188	1,082	996	913	858	818	783	732	684	639	594

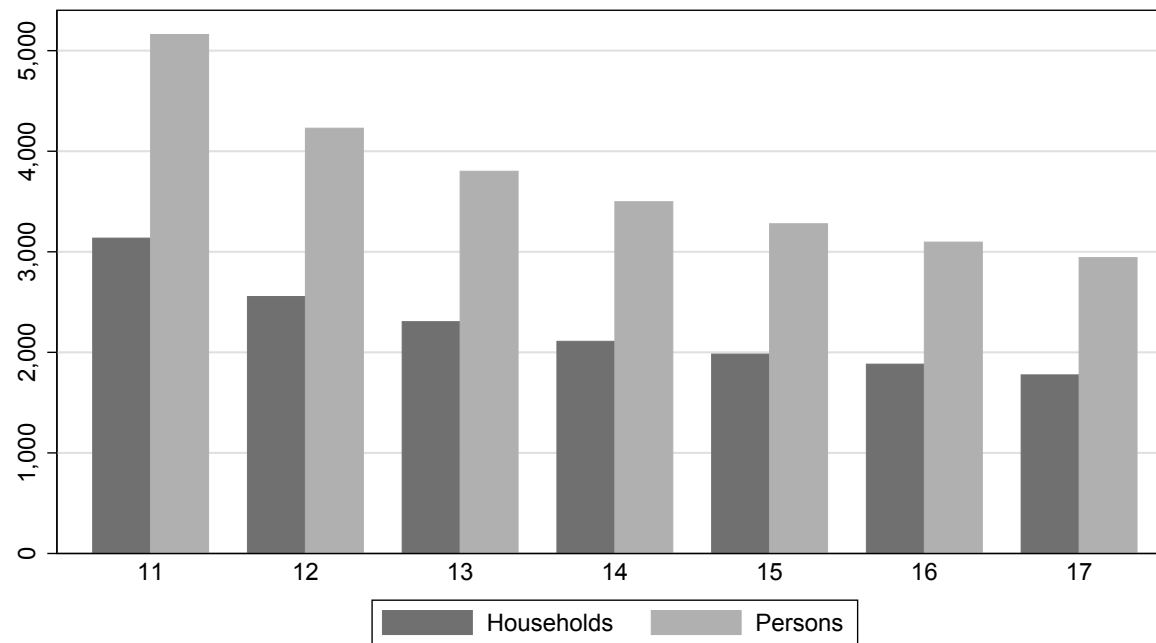


Figure 9: Comparison of Successful Interviews with Persons and Households (Subsample J), Waves 1 to 7

Year	2011	2012	2013	2014	2015	2016	2017
Persons	5,161	4,229	3,801	3,498	3,279	3,096	2,942
Households	3,136	2,555	2,305	2,110	1,983	1,883	1,776

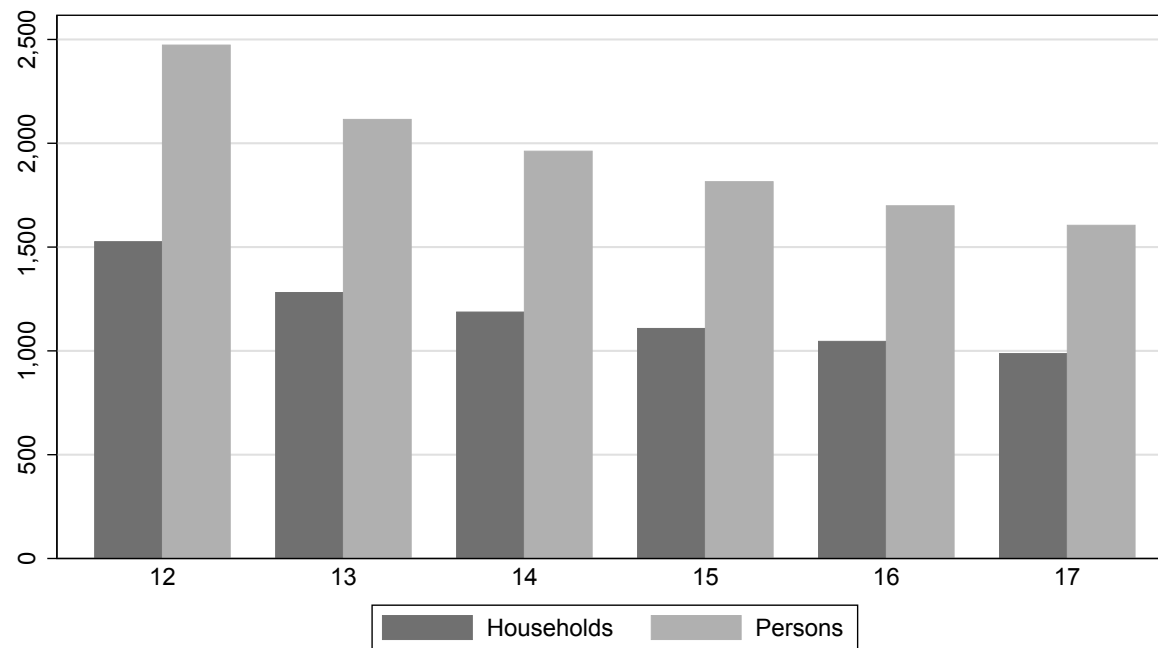


Figure 10: Comparison of Successful Interviews with Persons and Households (Subsample K), Waves 1 to 6

Year	2012	2013	2014	2015	2016	2017
Persons	2,473	2,115	1,962	1,815	1,699	1,605
Households	1,526	1,281	1,187	1,108	1,046	987

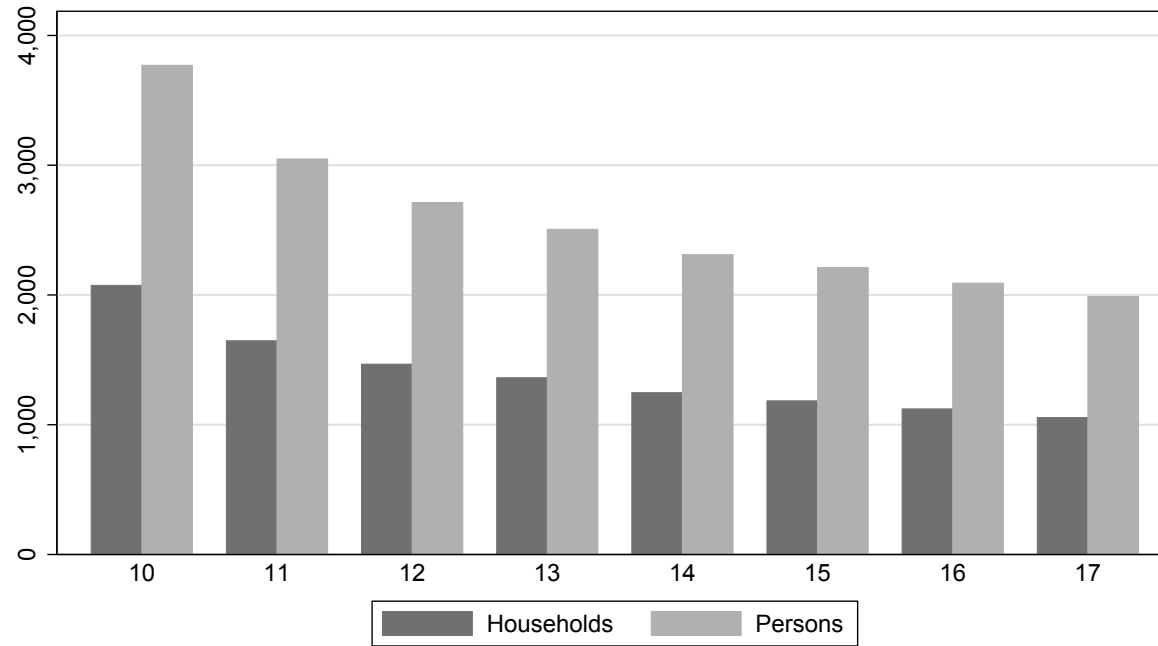


Figure 11: Comparison of Successful Interviews with Persons and Households (Subsample L1), Waves 1 to 8

Year	2010	2011	2012	2013	2014	2015	2016	2017
Persons	3,770	3,048	2,713	2,506	2,311	2,211	2,091	1,988
Households	2,074	1,647	1,467	1,362	1,247	1,184	1,122	1,055

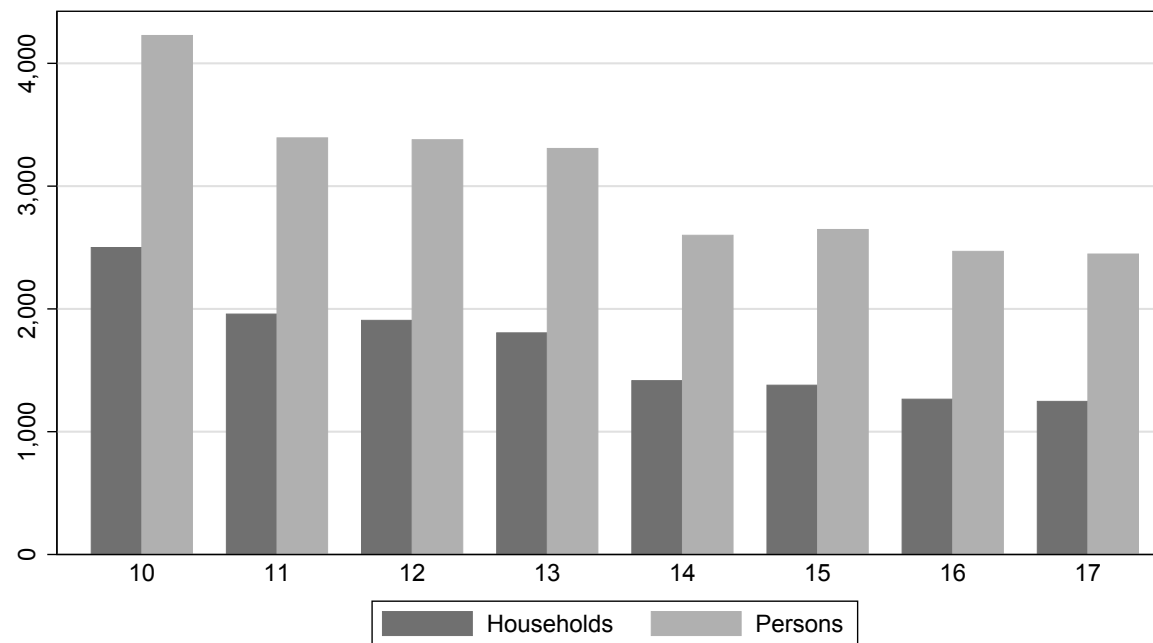


Figure 12: Comparison of Successful Interviews with Persons and Households (Subsample L2), Waves 1 to 8 ^{23,24}

Year	2010	2011	2012	2013	2014	2015	2016	2017
Persons	4,227	3,393	3,378	3,307	2,600	2,647	2,469	2,447
Households	2,500	1,958	1,907	1,805	1,416	1,379	1,265	1,247

²³237 households were identified not to be part of the target population and were not followed in the second wave.

²⁴In 2014 the default interview mode changed to Computer-Assisted Web Interviewing (CAWI).

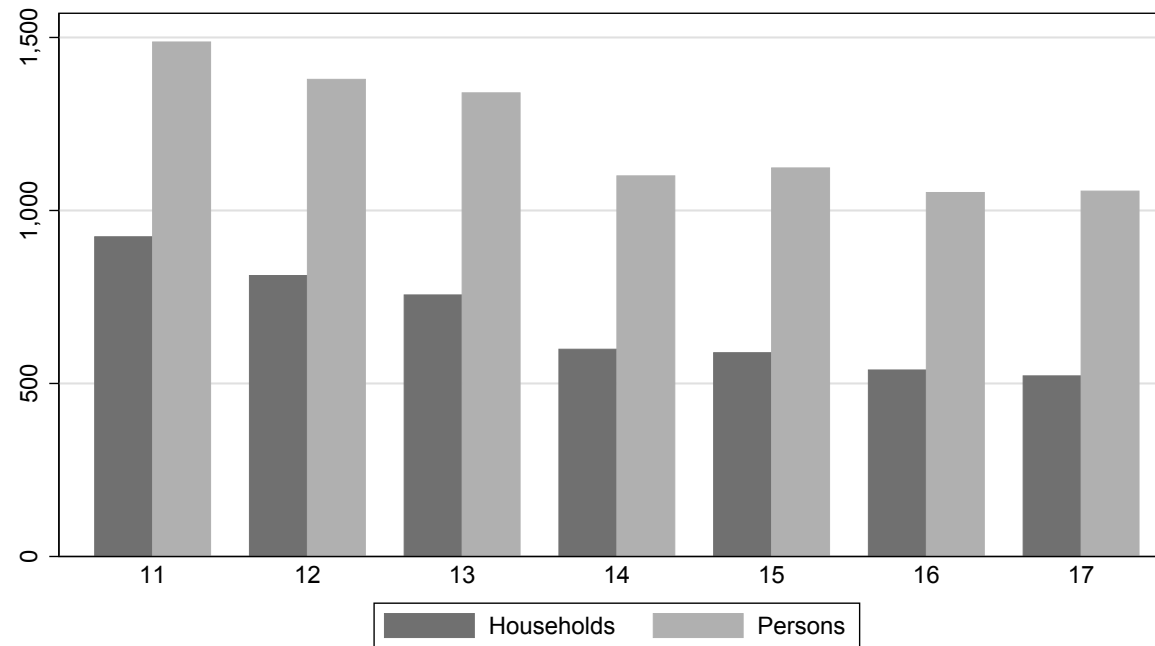


Figure 13: Comparison of Successful Interviews with Persons and Households (Subsample L3), Waves 1 to 7²⁵

Year	2011	2012	2013	2014	2015	2016	2017
Persons	1,487	1,379	1,340	1,100	1,123	1,052	1,056
Households	924	812	756	599	589	539	522

²⁵In 2014 the default interview mode changed to Computer-Assisted Web Interviewing (CAWI).

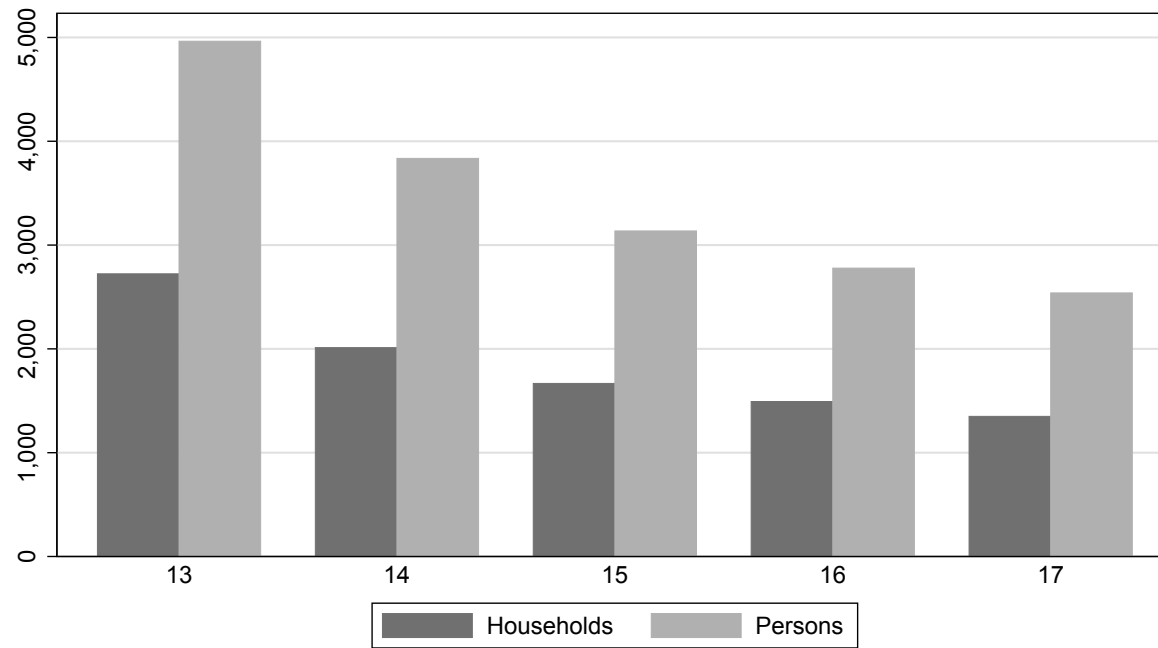


Figure 14: Comparison of Successful Interviews with Persons and Households (Subsample M1), Waves 1 to 5

Year	2013	2014	2015	2016	2017
Persons	4,964	3,835	3,136	2,778	2,539
Households	2,723	2,012	1,667	1,493	1,350

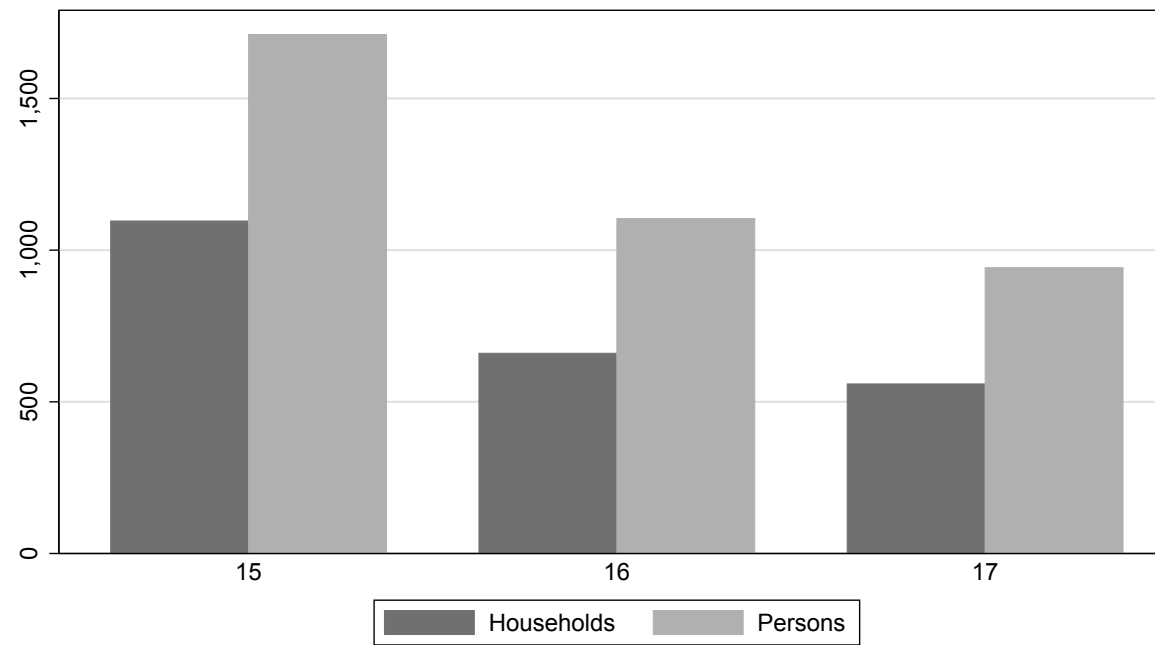


Figure 15: Comparison of Successful Interviews with Persons and Households (Subsample M2), Waves 1 to 3

Year	2015	2016	2017
Persons	1,711	1,104	942
Households	1,096	660	559

2.2 Continuance and Exit: The First Wave Gross Samples and their Participation 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”).

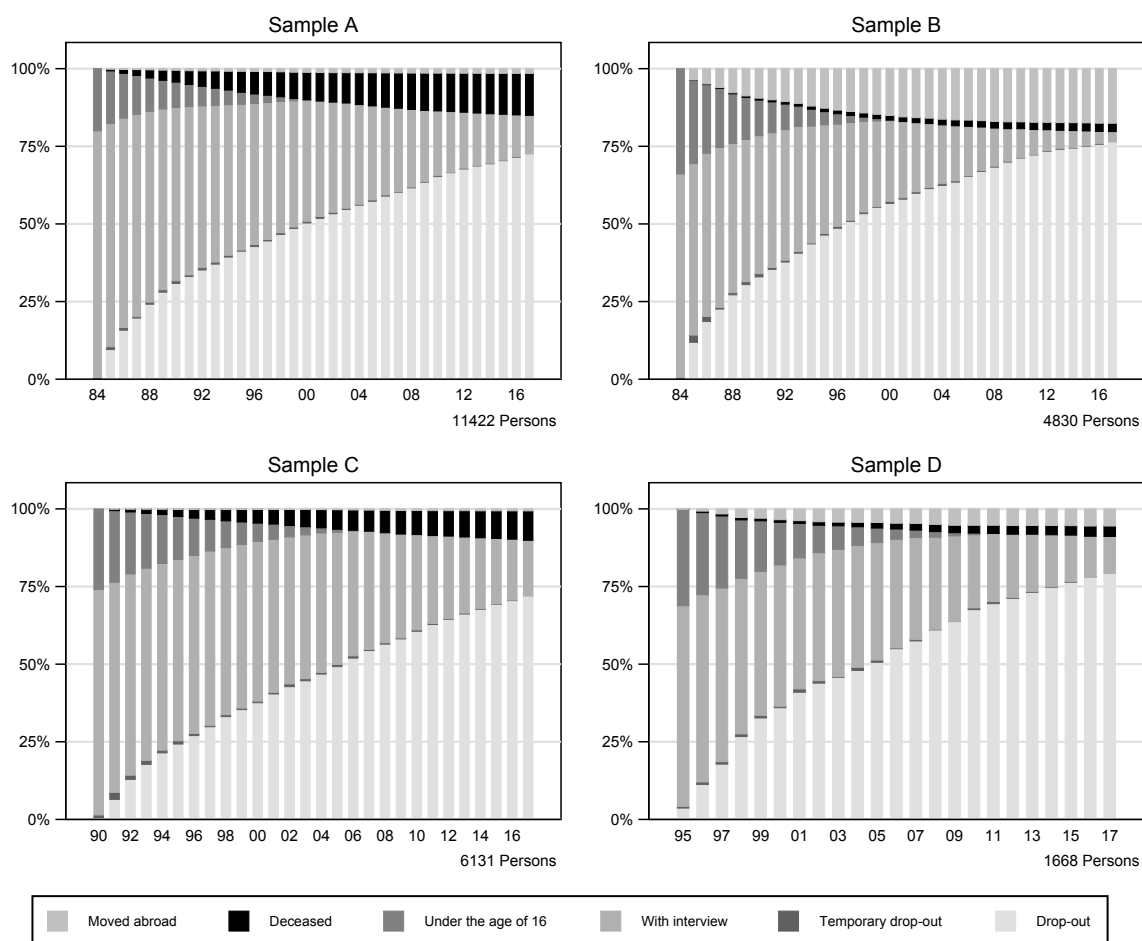


Figure 16: First-Wave Persons and their Participation Behavior. Development up to 2017

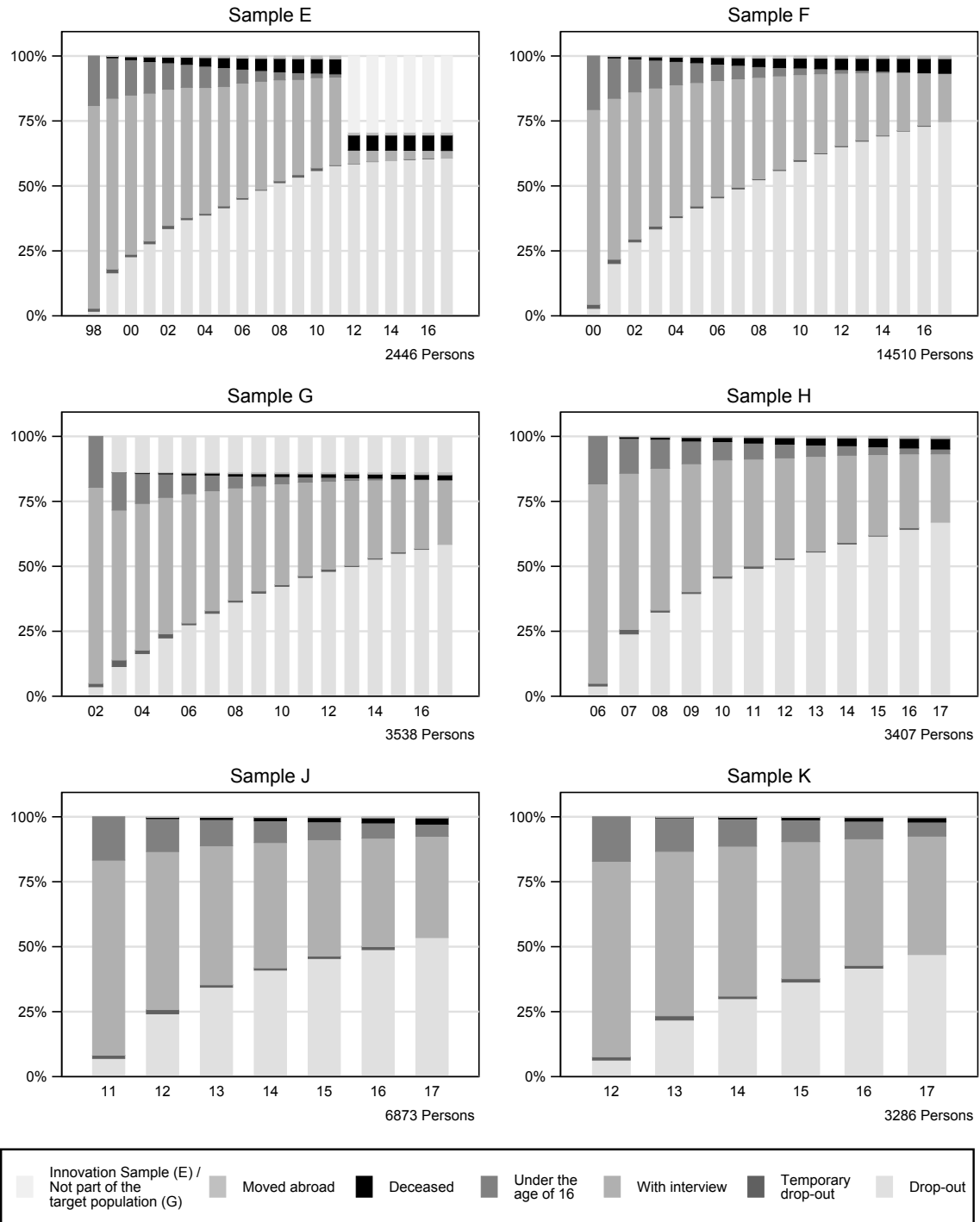


Figure 16: First-Wave Persons and their Participation Behavior. Development up to 2017

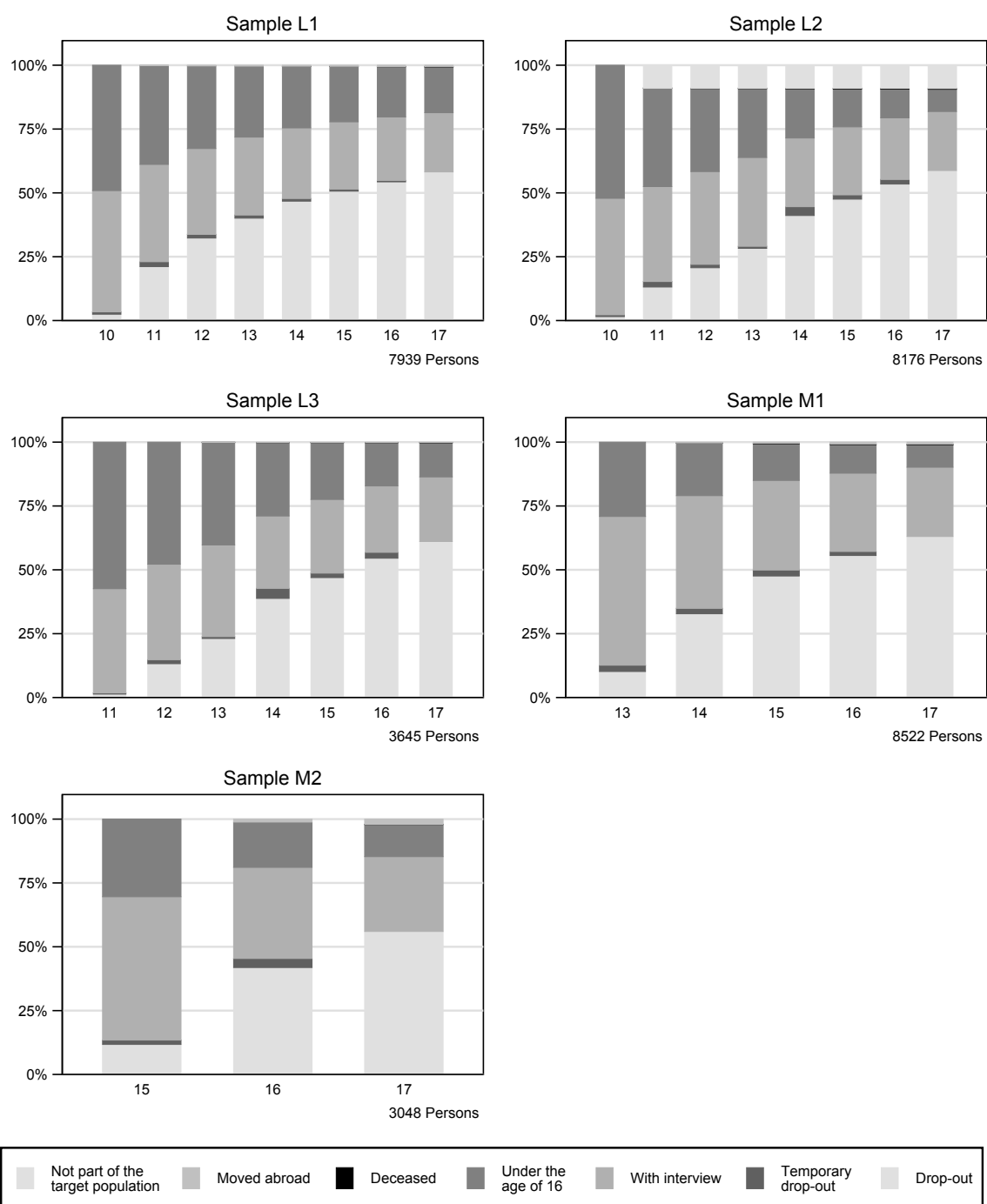


Figure 16: First-Wave Persons and their Participation Behavior. Development up to 2017

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.

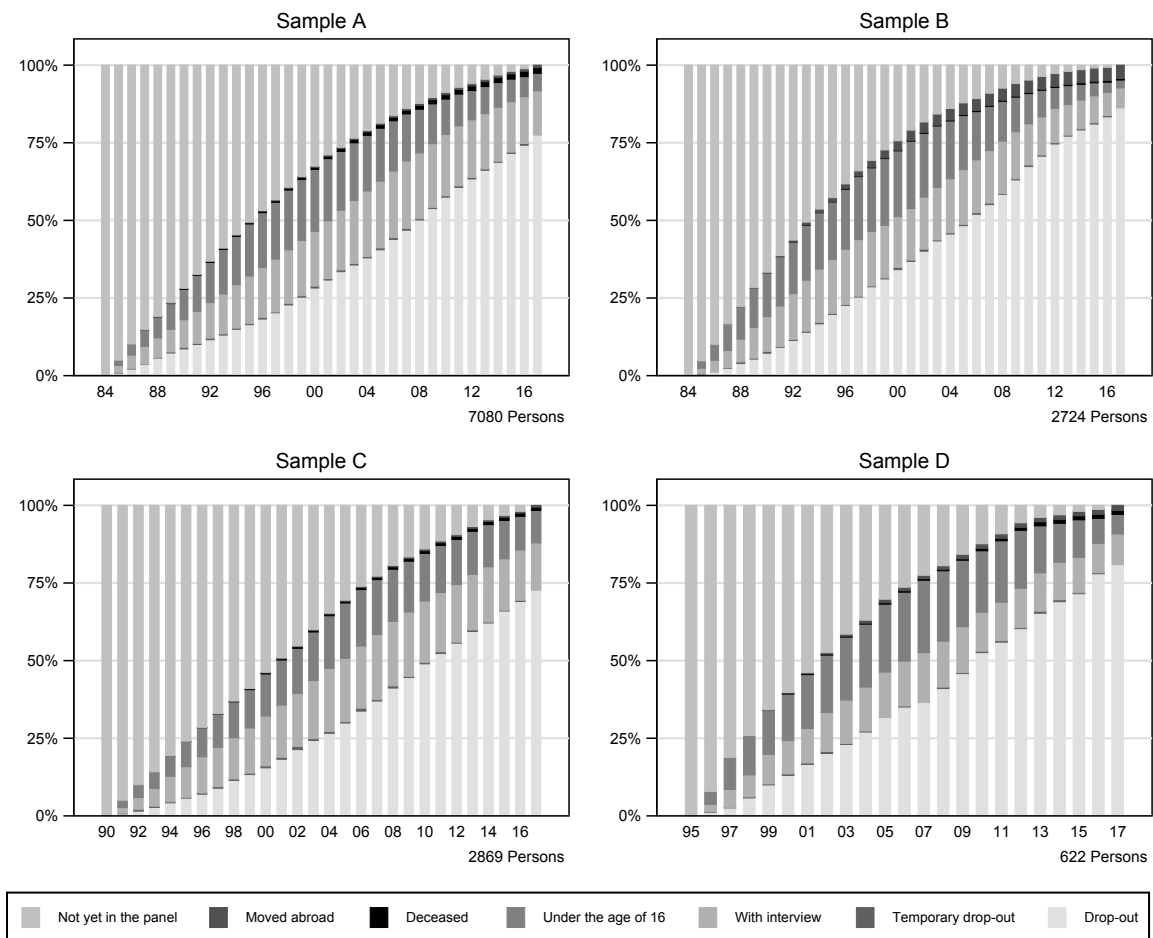


Figure 17: Entrants and their Participation Behavior. Development up to 2017

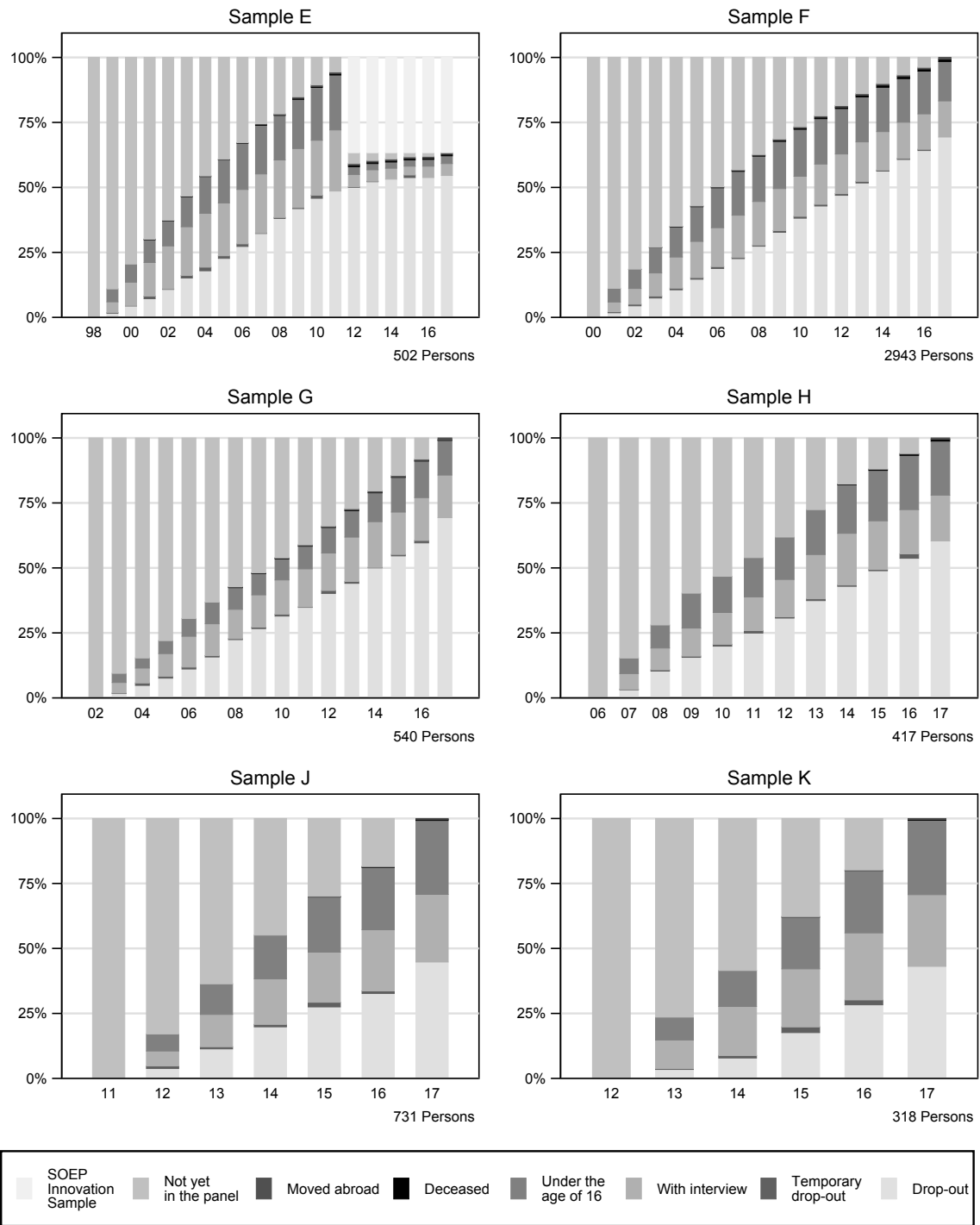


Figure 17: Entrants and their Participation Behavior. Development up to 2017

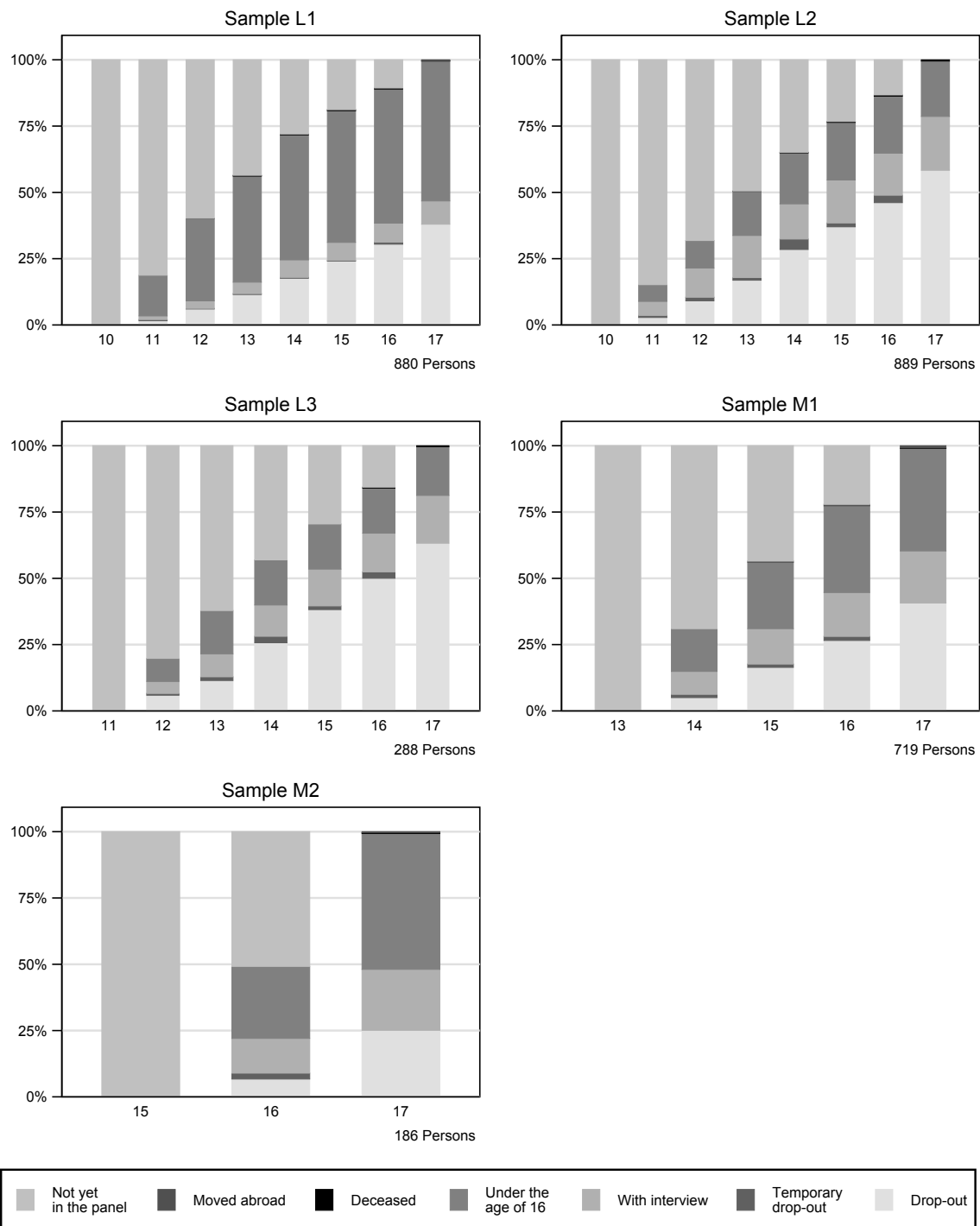


Figure 17: Entrants and their Participation Behavior. Development up to 2017

2.4 Original Households and Split-Offs

In case a household splits in multiple households (for instance, because a household member moves into another apartment), all resulting split-off households will be interviewed. The household which is not moving keeps the initial household number. These households are referred to as an “original household”²⁶. The following figures display the development of the share of original households for each sample.

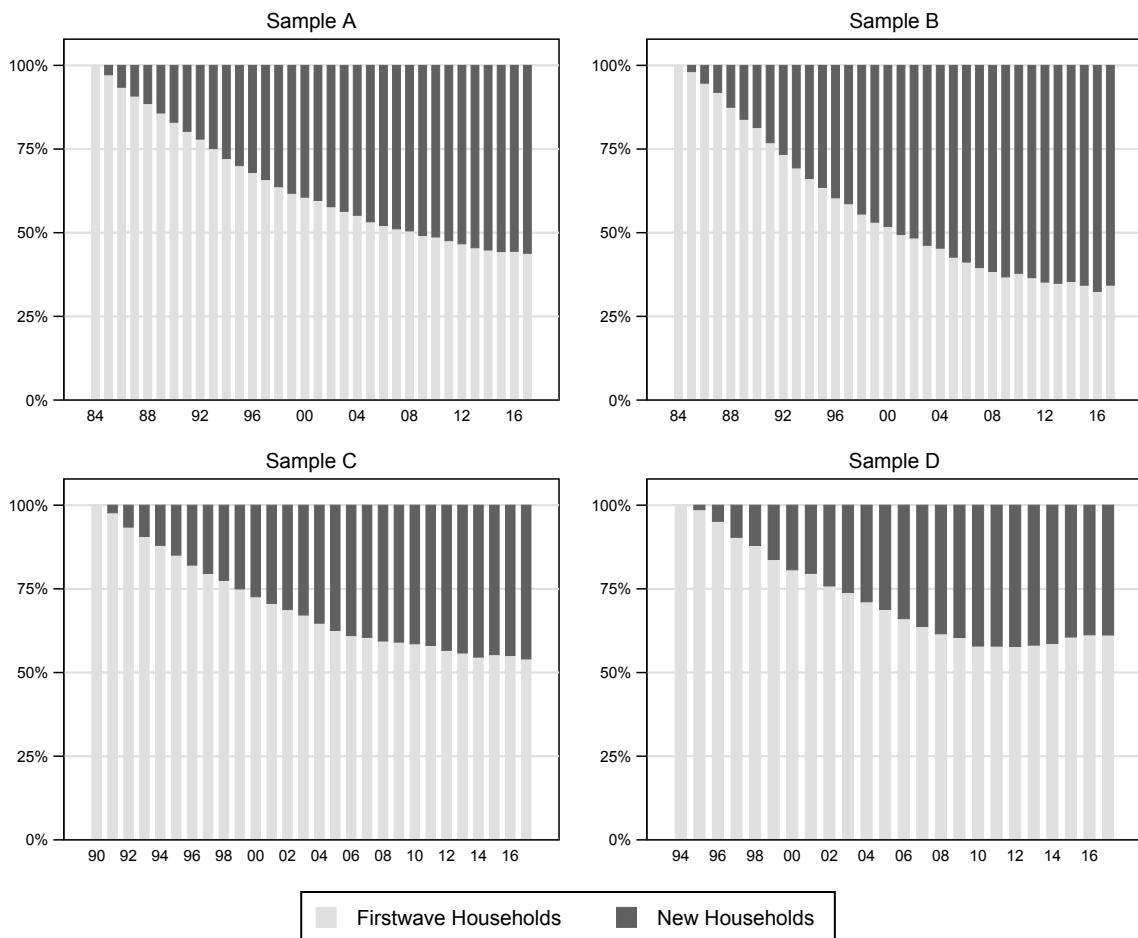


Figure 18: Proportion of First-Wave and New Households. Development up to 2017

²⁶For detailed studies on the relevance of non-original sample members in the SOEP, see Schonlau et al. (2011) and Spiess et al. (2008).

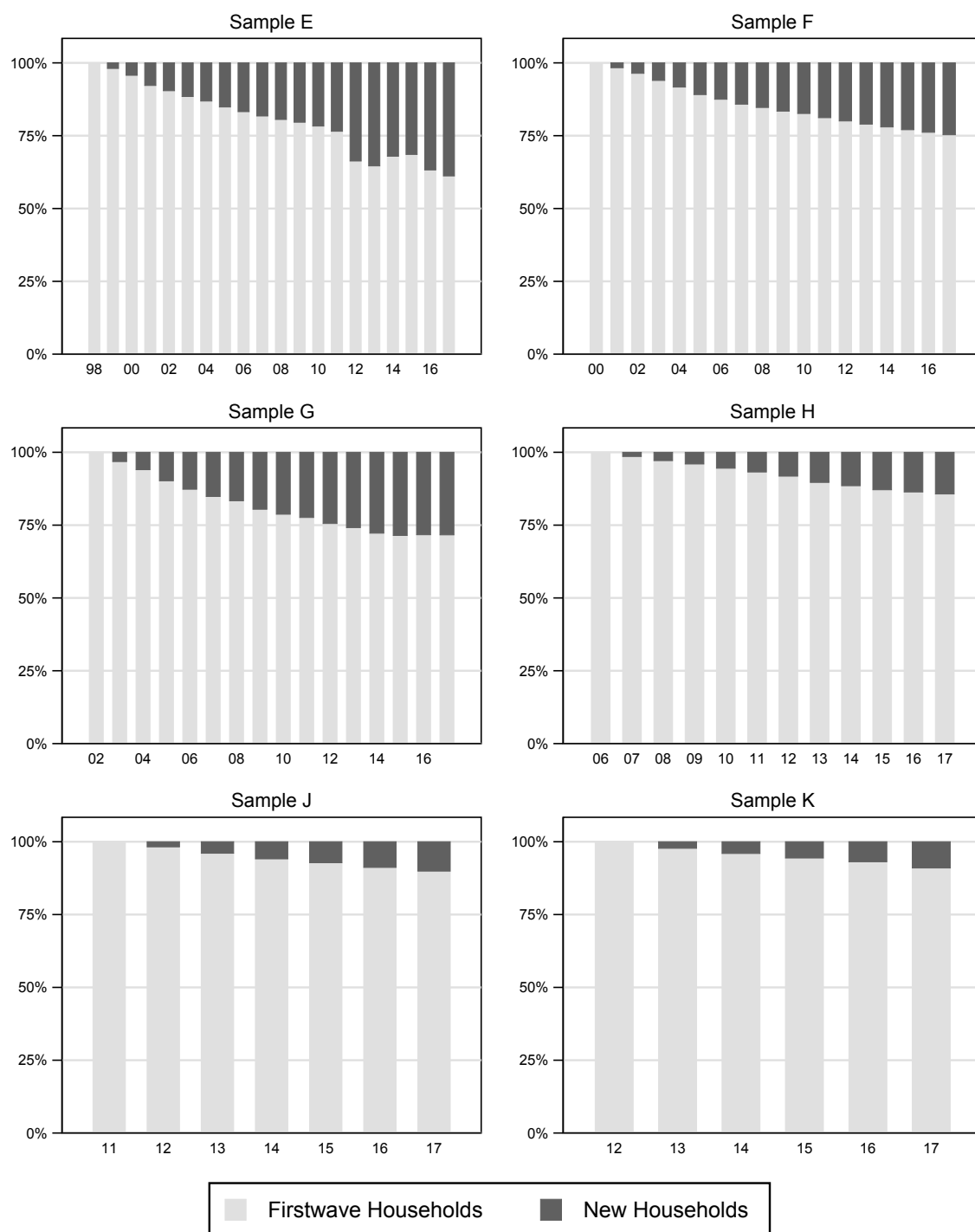


Figure 18: Proportion of First-Wave and New Households. Development up to 2017

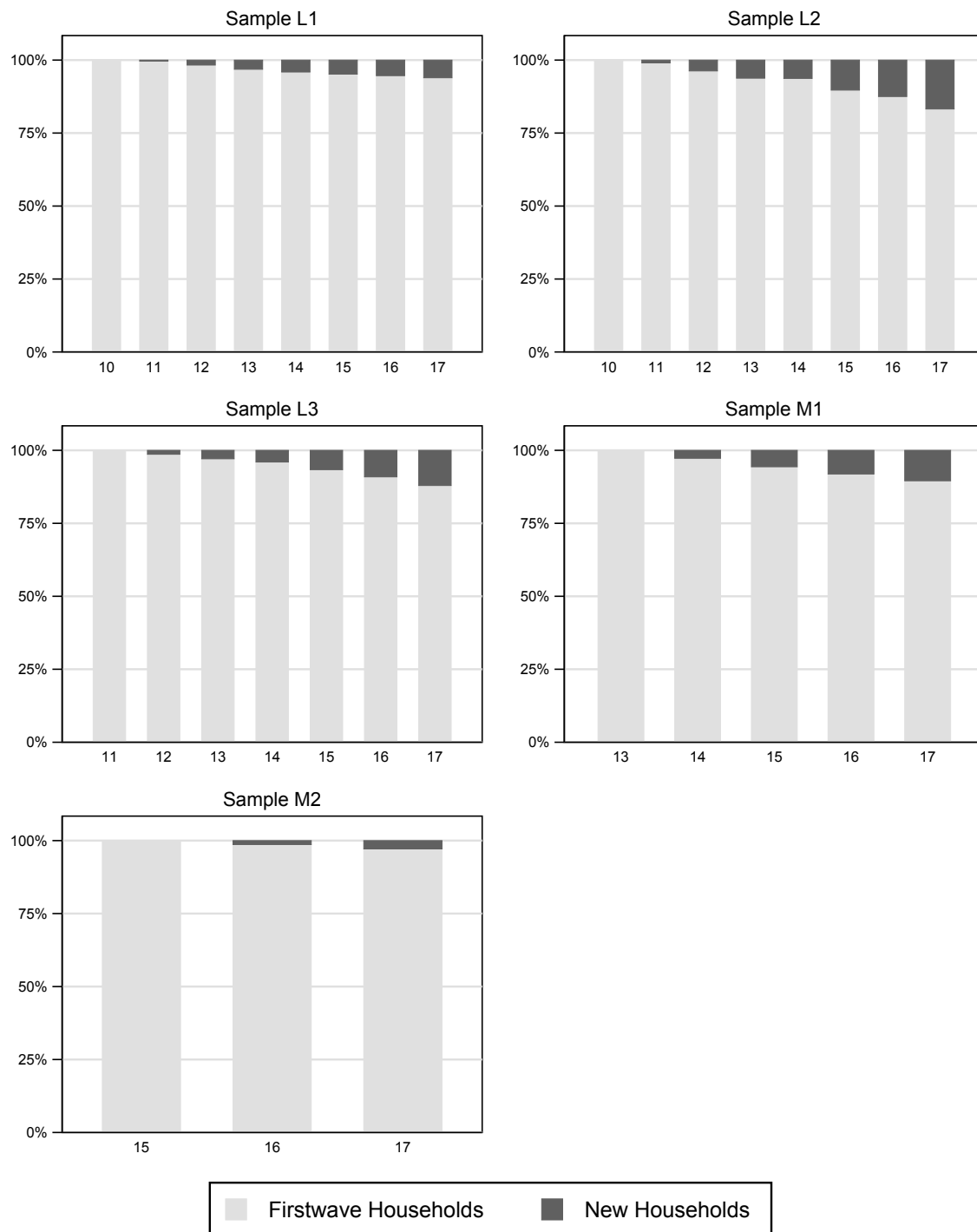


Figure 18: Proportion of First-Wave and New Households. Development up to 2017

2.5 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 19, 20, 21, 22, and 23) and some basic socio-demographic characteristics measured in the year of sampling, such as age, occupation, income, and education (Figures 24, 25, 26 and 27). 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 19), 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.

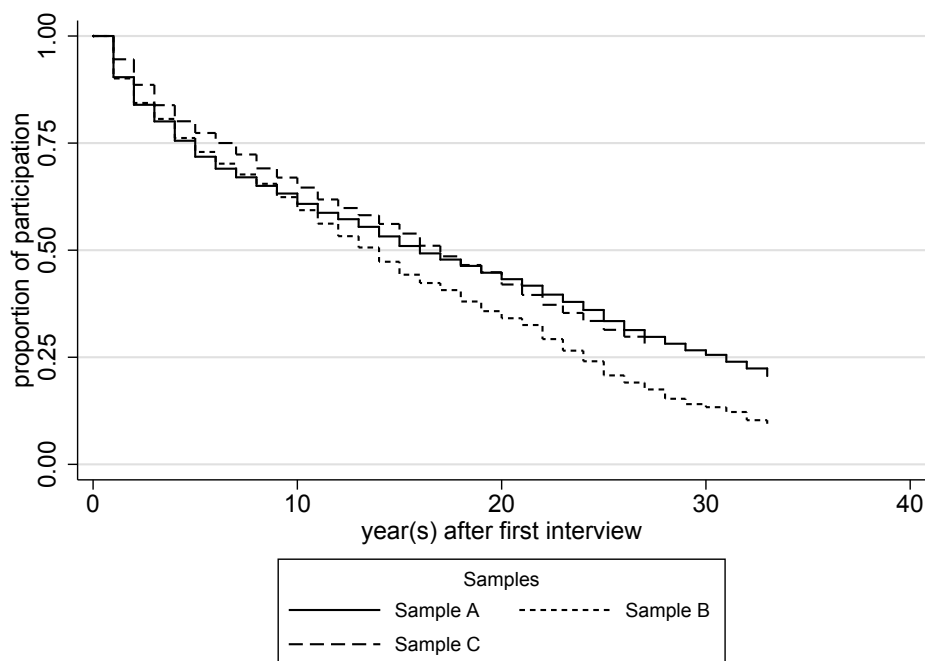


Figure 19: Successful Re-Interviewing of First-Wave Respondents by Subsamples A, B, C. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

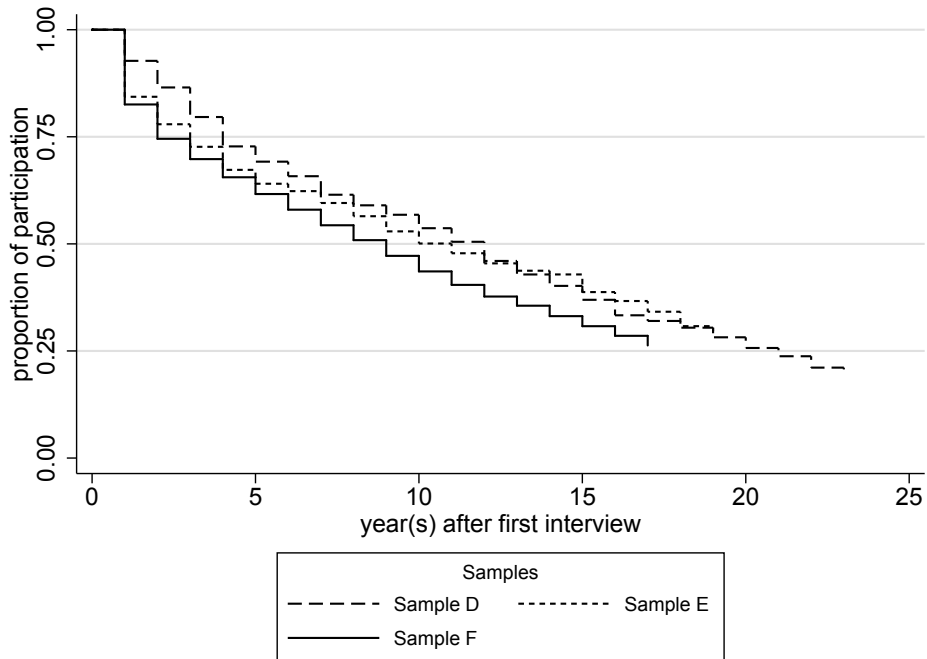


Figure 20: Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

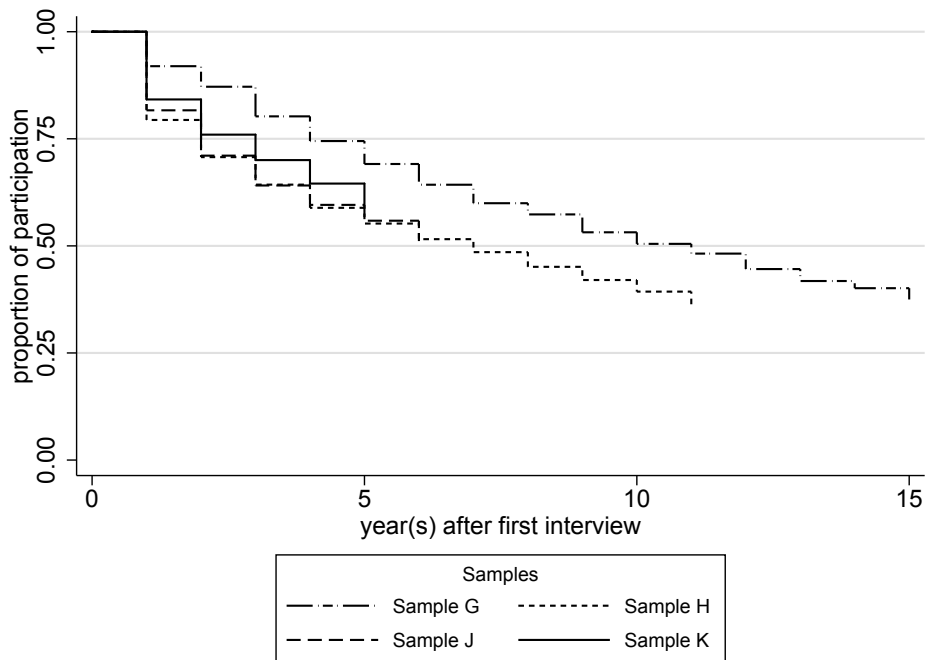


Figure 21: Successful Re-Interviewing of First-Wave Respondents by Subsamples G, H, J and K. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

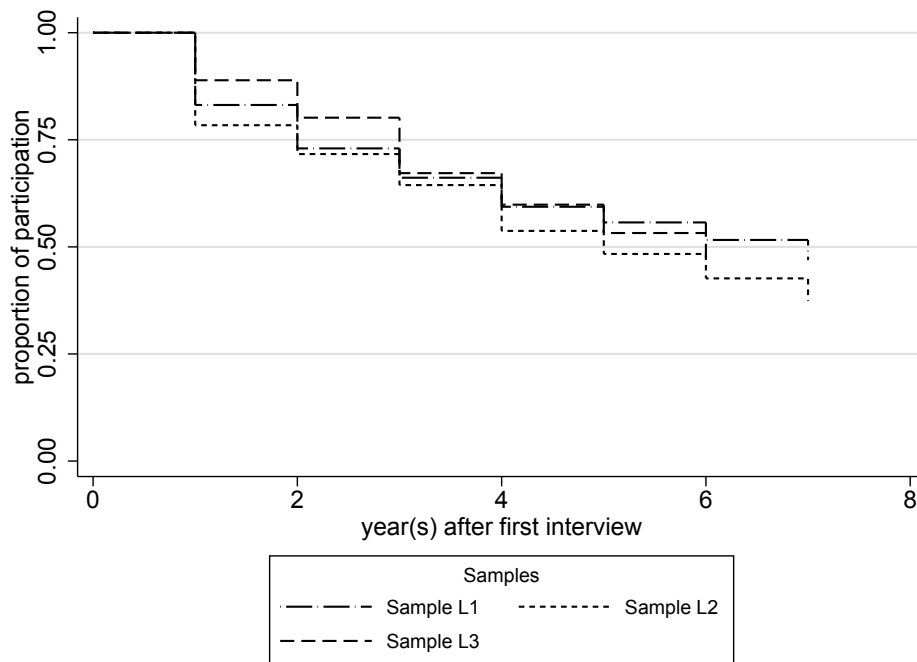


Figure 22: Successful Re-Interviewing of First-Wave Respondents by Subsamples L1, L2 and L3. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

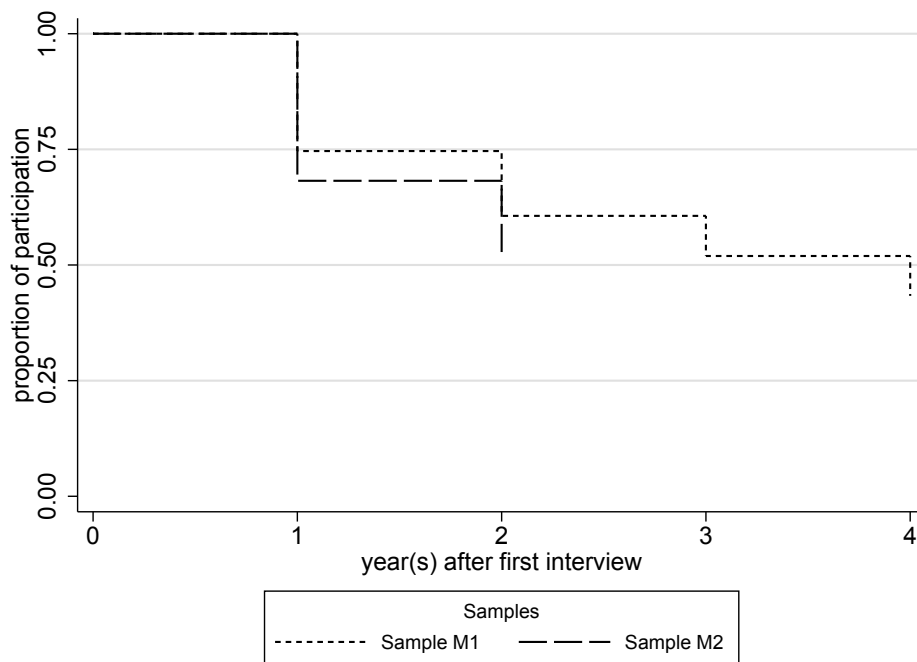


Figure 23: Successful Re-Interviewing of First-Wave Respondents by Subsamples M1 and M2. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

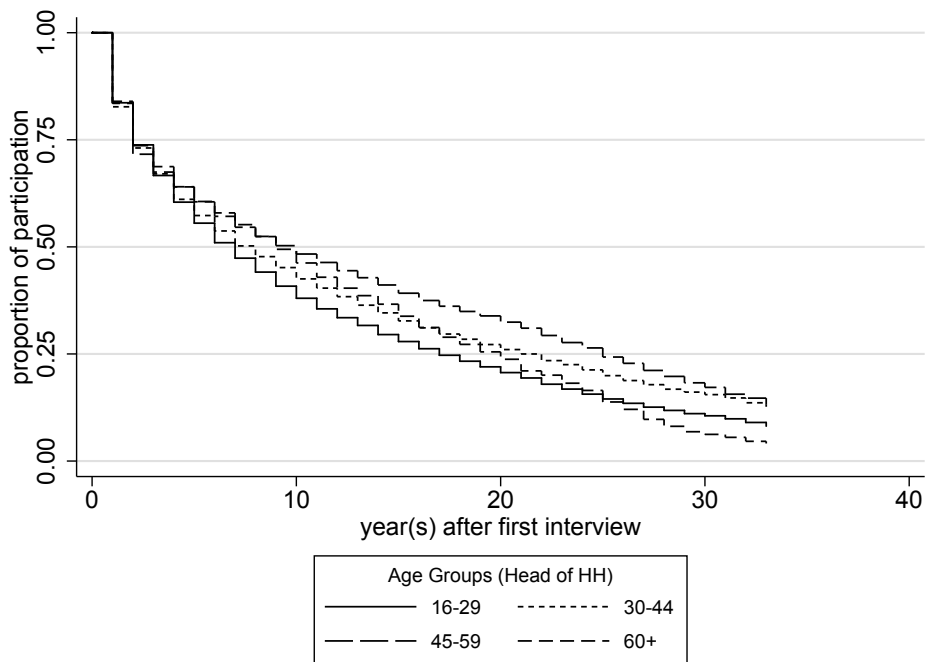


Figure 24: Successful Re-Interviewing of All First-Wave Respondents by Age Categories. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

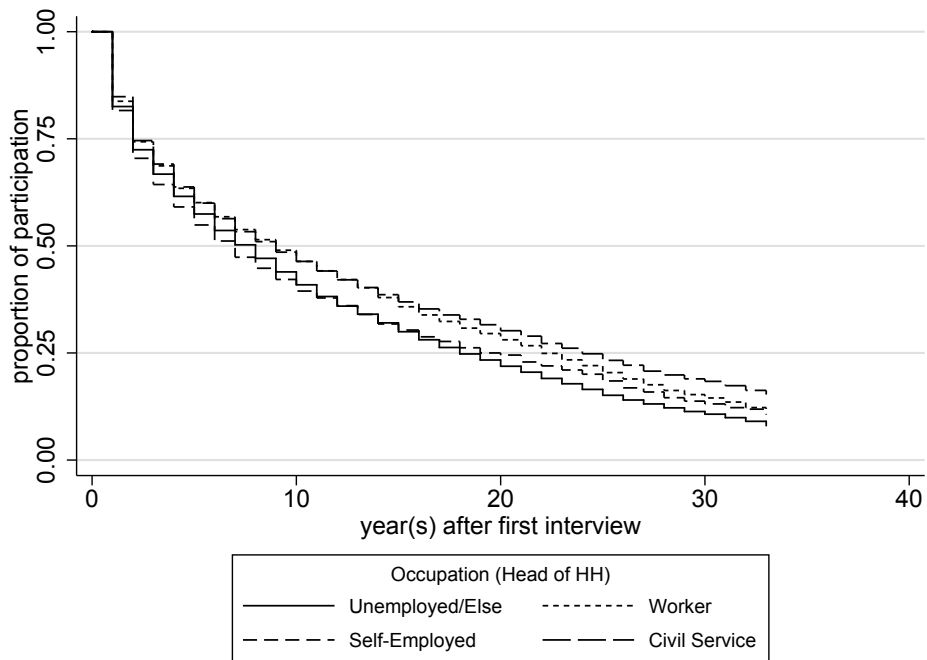


Figure 25: Successful Re-Interviewing of All First-Wave Respondents by Occupation. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

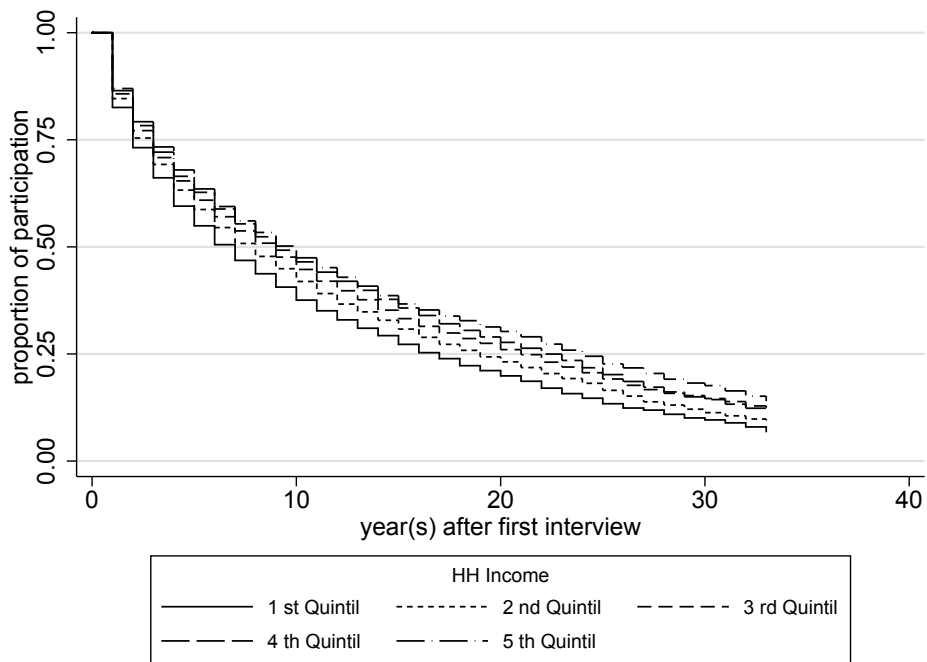


Figure 26: Successful Re-Interviewing of All First-Wave Respondents by Income Quintiles. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

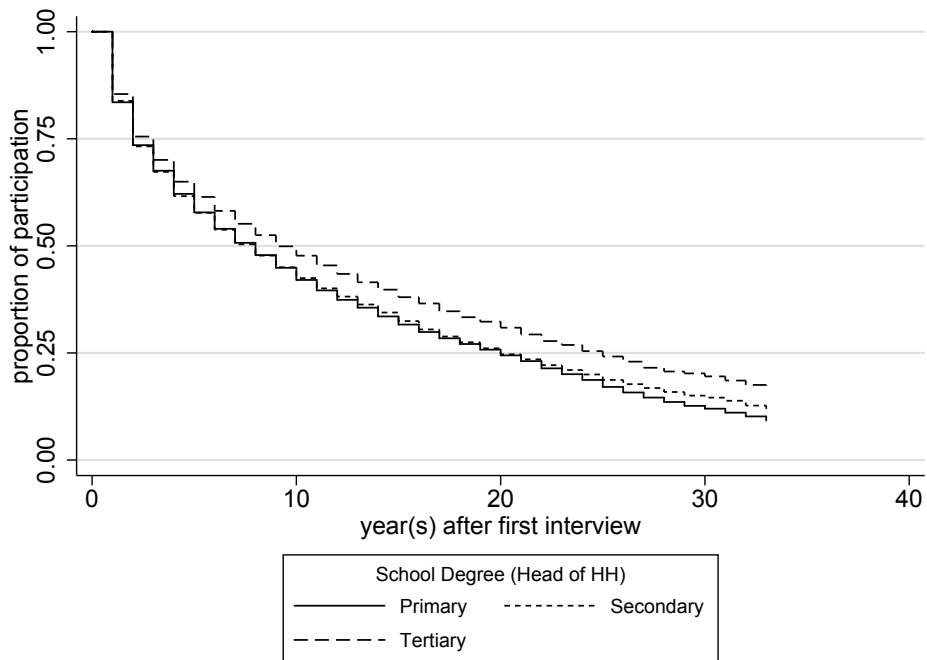


Figure 27: Successful Re-Interviewing of All First-Wave Respondents by Education. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

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, (c) all household members have left the sampling area or all household members have died, and (d) all household members have returned to an existing panel household.

3.1 The Frequency of Successful Follow-Ups

Table 3.1 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 M2 and waves 1985 through 2017. 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 N by Year.

Year	Sample A		Sample B		Sample C		Sample D		Sample E		Sample F		Sample G		Sample H		Sample I		Sample J		
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	
1984	4,528		1,393																		
1985	4,681	98.5	1,370	96.8																	
1986	4,486	99.0	1,325	97.4																	
1987	4,232	99.1	1,220	98.7																	
1988	4,140	99.2	1,191	99.1																	
1989	3,984	99.1	1,157	99.0																	
1990	3,902	99.2	1,124	98.8	2,179																
1991	3,860	99.5	1,151	99.3	2,246	98.5															
1992	3,845	99.7	1,153	99.2	2,302	99.5															
1993	3,867	99.3	1,172	98.6	2,227	99.1															
1994	3,849	99.3	1,150	99.0	2,134	99.4	236														
1995	3,784	99.5	1,108	99.0	2,110	99.6	540	100.0													
1996	3,747	99.7	1,069	99.3	2,103	99.5	544	99.6													
1997	3,688	99.6	1,038	99.1	2,087	99.5	541	99.3													
1998	3,667	99.4	1,019	99.4	2,079	99.4	528	99.1	1,056												
1999	3,631	99.6	975	99.4	2,037	99.7	498	99.4	1,089	99.5											
2000	3,549	99.6	934	99.5	2,025	99.7	467	99.8	967	99.2	6,043										
2001	3,463	99.6	904	99.4	2,034	99.7	454	99.1	921	99.1	6,162	99.0									
2002	3,406	99.7	877	99.1	2,005	99.6	450	99.8	873	99.4	5,447	99.5	1,224								
2003	3,330	99.6	840	99.6	1,982	99.6	434	99.5	834	99.3	4,965	99.7	1,056	99.1							
2004	3,260	99.8	803	99.6	1,962	99.6	436	99.8	797	99.7	4,736	99.6	1,010	99.7							
2005	3,220	99.8	779	99.4	1,959	99.7	429	99.3	783	99.9	4,577	99.7	1,001	99.7							
2006	3,138	99.7	770	99.6	1,941	99.4	425	98.8	775	99.1	4,401	99.3	995	99.5	1,506						
2007	3,000	99.7	725	99.4	1,834	99.9	387	99.5	727	99.7	4,157	99.5	933	99.2	1,530	99.5					
2008	2,856	99.8	676	99.3	1,767	99.5	372	99.5	680	99.7	3,962	99.4	904	99.7	1,326	99.6					
2009	2,730	99.7	620	99.4	1,695	99.9	351	99.7	636	100.0	3,760	99.6	870	99.5	1,145	99.7	1,495				
2010	2,570	99.8	548	99.5	1,627	100.0	334	99.7	605	99.8	3,538	99.6	826	99.9	1,059	99.5	1,738	98.3			
2011	2,421	99.8	495	99.2	1,541	99.8	303	99.3	589	100.0	3,319	99.7	797	99.6	992	99.6				3,136	
2012	2,289	99.8	440	99.8	1,466	99.9	286	100.0	116	99.1	3,076	99.9	774	99.7	928	99.9				3,204	99.2
2013	2,180	99.6	393	99.2	1,417	99.7	269	99.3	98	100.0	2,881	99.7	733	99.6	877	99.5				2,871	99.5
2014	2,078	99.4	361	99.4	1,351	99.6	249	100.0	90	100.0	2,741	99.7	725	99.3	828	99.4				2,519	99.1
2015	1,998	99.4	331	99.4	1,300	99.5	229	100.0	83	100.0	2,597	99.2	699	99.3	790	99.7				2,309	99.4
2016	1,861	99.6	296	99.7	1,217	99.7	208	99.5	83	96.4	2,412	99.4	669	98.7	720	99.6				2,119	99.5
2017	1,748	99.3	271	98.9	1,125	99.6	184	99.5	75	100.0	2,214	99.3	622	99.5	677	99.4				2,015	99.3

Year	Sample K		Sample L1		Sample L2		Sample L3		Sample M1		Sample M2		Sample M3/4		Sample M5		Sample N					
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%				
2010			2,074		2,500																	
2011			2,083	98.8	2,271	98.1	924															
2012	1,526		1,867	99.6	2,255	98.5	943	98.7														
2013	1,564	99.0	1,753	99.3	2,177	98.8	920	99.1	2,723													
2014	1,448	99.4	1,512	99.4	2,027	98.2	836	98.6	2,828	98.8												
2015	1,308	99.3	1,404	99.4	1,880	98.4	789	97.8	2,456	98.0	1,096											
2016	1,209	99.3	1,287	99.5	1,736	98.6	732	98.2	2,116	97.7	1,096	97.1	3,289									
2017	1,105	99.5	1,209	99.3	1,587	98.6	686	97.8	1,794	97.9	931	98.1	3,351 ²⁷	93.6	1,519					2,314		

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

²⁷This number contains 112 cases that had to be deleted in 2016, due to incorrectly conducted interviews, and that were subsequently surveyed in 2017. Furthermore 112 cases had to be deleted in 2017 due to incorrectly conducted interviews.

3.2 Predicting the Probability of Successful vs. Unsuccessful Follow-Ups in the Year 2017

Based on household and interview level characteristics measured in 2016, we aim at predicting the probability of re-contacting a household relative to unsuccessful follow-up in 2017. 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 of the previous waves from 1985 to 2016 are not reported in the present data documentation due to space restrictions. 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 Household	Household new in SOEP	0/1
New Address	Household moved	0/1
New SOEP Member	Head of household had less than 4 interviews	0/1
Temporary Drop-Out	Temporary drop-out of household in prev. year	0/1
Drop-Out Related Household	Ultimate drop-out of related household	0/1
Change of Interviewer	Change of interviewer between the two last waves	0/1
Interviewer Related HH	Same interviewer in related household	0/1
Long Interview	Duration of the interview more than 20 Minutes	0/1
Phone Unknown	Telephone number undisclosed	0/1
Subsample M4	Household from Sample M4, Refugee/Families	0/1
Demographic Characteristics		
Non-German Nationality	Household member(s) has nationality other than German	0/1
Single Household	One-person household	0/1
Sibling in HH	Head of household has sibling(s) living in same household	0/1
Child Under 12	At least one child younger than 12 years in household	0/1
Work and Economic Situation		
Low Income	Low household income, within the 1st quartile	0/1
Low Expenditure	Household expenditures lower than income	0/1
Asylum Benefits Low	Head of household: Amount of asylum seeker benefits (AsylbLG) last month - 1st quartile	0/1
Same Employer 4th Q.	Head of household: Duration working for the current employer, 4th Q.	0/1
Telephone	Household has one or more telephone landlines	0/1
No Shoes	Not everyone in the household has at least two pairs of outdoor shoes	0/1
Personality and Attitudes		
Outsider Feeling	Head of household sometimes/often feels like an outsider	0/1
Positive Self-Attitude	Household member(s) has positive attitude toward themselves	0/1
Extreme Answers	High number of extreme answers of attitude questions 4th Q.	0/1
Building, Area, and Region		
Bavaria	Household located in Bavaria	0/1
Hesse	Household located in Hesse	0/1
High-Rise Building	Household lives in a high-rise building (9 or more stories)	0/1
High-Rise Area	Household located in area with many high-rise buildings	0/1
Small Apartments Area	Household located in area with high share of small apartments	0/1
High Average Age Area	Household located in area with high average age	0/1

Table 3.3: Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2017

	Sample A	Sample C	Sample F	Sample H	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1	Sample M2	Sample M3/4
Intercept	2.25***	1.84***	1.91***	2.12***	2.30***	1.92***	1.83***	2.49***	2.12***	1.74***	2.17***	1.93***
Interview Characteristics												
New Household	-1.27***	-1.20***			-1.07***	-0.95***	-0.95***	-0.86**	-0.60*	-1.57***	-1.43***	-0.72***
New Address	-0.90***		-0.90***					-1.14***	-0.76*	-1.26***	-1.02***	
New SOEP member						-0.75*						
Temporary Drop-Out								-0.97***				
Drop-Out Related Household				-1.37**								
Change of Interviewer												-0.30***
Interviewer Related HH										0.83*		
Long Interview												0.15*
Phone Unknown					-0.68**		-0.63**	-0.39*		-0.98***		-0.57***
Subsample M4												0.22***
Demographic Characteristics												
Non-German Nationality								-0.57*				
Single Household									-1.02***	-0.62***	-0.69**	-0.31***
Sibling in HH								-0.72**				
Child Under 12												-0.23**
Work and Economic Situation												
Low Income												-0.20**
Low Expenditure												-0.39*
Asylum Benefits Low												-0.31***
Same employer 4th Q.				-0.90*	-0.60**							
Telephone												0.20*
No Shoes												-0.20**
Personality and Attitudes												
Outsider Feeling												-0.16**
Positive Self-Attitude										-0.88**		
Extreme Answers										0.74***		

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

In Samples B, D, E and G less than 4 households were not re-contacted.

Table 3.3 – *Continued from previous page*

	Sample A	Sample C	Sample F	Sample H	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1	Sample M2	Sample M3/4
Building, Area, and Region												
Bavaria												-0.17*
Hesse												0.52**
High-Rise Building									-0.64*			
High-Rise Area					-0.55**							
Small Apartments Area	-0.45*											
High Average Age Area			-0.28*									
<i>No. of Observations</i>	1,748	1,125	2,214	677	2,015	1,105	1,209	1,587	686	1,794	931	3,238
<i>Log Likelihood</i>	-43.73	-23.08	-67.29	-15.09	-43.69	-23.55	-36.40	-73.36	-45.30	-90.26	-55.09	-617.22

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

In Samples B, D, E and G less than 4 households were not 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 subsample 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 N by Year.

Year	Sample A		Sample B		Sample C		Sample D		Sample E		Sample F		Sample G		Sample H		Sample I		Sample J		
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	
1984	4,528		1,393																		
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																
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														
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												
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										
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								
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						
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				
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	1,709	68.8			
2011	2,417	88.9	491	79.6	1,538	88.1	301	88.4	589	92.5	3,308	87.2	794	88.9	988	86.8				3,136	
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				3,179	80.4
2013	2,172	89.7	390	82.3	1,413	88.5	267	86.9	98	83.7	2,878	89.3	730	92.7	873	89.7				2,857	80.7
2014	2,065	90.8	359	84.1	1,346	90.0	249	85.5	90	86.7	2,732	88.4	720	89.0	823	88.9				2,497	84.5
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
2016	1,853	87.9	295	77.3	1,213	88.5	207	83.6	80	85.0	2,398	87.3	660	89.4	717	89.1				2,108	89.3
2017	1,736	88.0	268	75.0	1,120	89.0	183	90.2	75	89.3	2,199	89.5	619	90.6	673	88.3				2,001	88.8

Year	Sample K		Sample L1		Sample L2		Sample L3		Sample M1		Sample M2		Sample M3/4		Sample M5		Sample N				
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%			
2010			2,074		2,500																
2011			2,057	80.1	2,228	87.9	924														
2012	1,526		1,859	78.9	2,222	85.9	931	87.2													
2013	1,549	82.7	1,741	78.2	2,151	83.9	912	82.9	2,723												
2014	1,439	82.5	1,503	83.0	1,990	71.2	824	72.7	2,793	72.0											
2015	1,299	85.3	1,396	84.8	1,850	74.5	772	76.3	2,407	69.3	1,096										
2016	1,201	87.1	1,280	87.7	1,712	73.9	719	75.0	2,067	72.2	1,054	62.0	3,289								
2017	1,099	89.8	1,200	87.9	1,564	79.7	671	77.8	1,757	76.8	913	61.2	3,138 ²⁸	73.0	1,519					2,314	

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

²⁸This number contains 112 cases that had to be deleted in 2016, due to incorrectly conducted interviews, and that were subsequently surveyed in 2017. Furthermore 112 cases had to be deleted in 2017 due to incorrectly conducted interviews.

4.2 Predicting the Probability of Re-Interviewing vs. Refusal in the Year 2017

Based on the household and interview characteristics measured in the year 2016, and some regional information measured in 2017, we aim at predicting the probability of agreement vs. refusal to participate in the survey for households that were re-contacted in 2017. 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 2016 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		
CAPI	Computer Assisted Personal Interview	0/1
Email Known	Email address disclosed	0/1
Interviewer Related HH	Same interviewer in related household	0/1
Late Interview	Interview done in later months	0/1
Long Interview	Duration of the Interview more than 20 Minutes	0/1
Mode Change	Change of interview mode between the last two waves	0/1
New Address	Household moved	0/1
Mother-Child-Questionnaire	Additional mother-child-questionnaire in household	0/1
New Household	Household new in SOEP	0/1
Not Original Sample Member	Head of household is not an original sample member	0/1
High Item Nonresponse P.	Household member(s) have high item nonresponse in person questionnaire	0/1
Part. Unit Nonresponse	Household member(s) did not participate last wave	0/1
Unrealistic Answers	Interviewer considers participants answers rather unauthentic	0/1
Interview Related HH	Successful interview of related household	0/1
Temp. Drop-Out Related HH	Temporary drop-out of related household	0/1
Drop-Out of Related HH	Ultimate drop-out of related household	0/1
Multiple Children	Household from Sample L2/L3, Families with multiple children	0/1
Phone Unknown	Telephone number undisclosed	0/1
Temporary Drop-Out	Temporary drop-out of household in prev. year	0/1
Change of Interviewer	Change of interviewer between the two last waves	0/1
SOEP Household Moved	Participating household moved to new location	0/1
Subsample M4	Household from Sample M4, Refugee/Families	0/1
Demographic Characteristics		
Age 35-44	Head of household aged between 35 and 44 years	0/1
Age 65-74	Head of household aged between 65 and 74 years	0/1
Parents Abroad	Head of household: Both parents live abroad	0/1
Parents in HH	Head of household: Both parents live in household	0/1
Family Household	4 or more persons live in household	0/1
Partner not in HH	Partner of head of household does not live in the same household	0/1
Female Head of HH	Head of household is female	0/1
Single Household	One-person household	0/1

Table 4.2 – Continued from previous page

Variable	Label	Value
Child under 12	At least one child younger than 12 years in household	0/1
Work, Education and Financial Situation		
Same Employer 4th Q.	Head of household: Duration working for the current employer, 4th quartile	0/1
Receiving ALG II benefits	Household member(s) receives ALG II unemployment benefits	0/1
High Number of Books	Number of books in household: 200 or more	0/1
Low Number of Books	Number of books in household: below 50	0/1
House Owner	Head of household is owner of dwelling	0/1
No Car in HH	Household does not own a car	0/1
Retirement Concerns	Household member(s) worried about their own retirement pension	0/1
No Retirement Concerns	Household member(s) not worried about their own retirement pension	0/1
Open-Ended Contract	Head of household has a permanent employment contract	0/1
Interrupted at Work	Household member(s) often interrupted/bothered at work	0/1
Temporary Work	Head of household in temporary or contract employment	0/1
Job Intention	Head of household intends to obtain employment in the future	0/1
Limited Illegal Work	Household member(s): Only little illegal work in personal environment	0/1
Health, Well-Being and Personality Traits		
No Health Restrictions	Household member(s) not restricted in daily life due to health problems	0/1
Angry	Head of household often felt angry in the last 4 weeks	0/1
Calm and Relaxed	Head of household often felt calm and relaxed in the last 4 weeks	0/1
Unhappy	Head of household often felt unhappy in the last 4 weeks	0/1
No Health Concerns	Head of household is not concerned about own health	0/1
Chronic Back Pain	Household member(s) diagnosed with chronic back pain	0/1
Overweight/High BMI	Household member(s) has a body mass index of 30 or higher	0/1
Free Speech as Goal	Household member(s) considers free speech very important goal of politics	0/1
Often Eats Poultry	Household member(s) very often eats poultry	0/1
Never Eat Fish	Household member(s) never eat fish	0/1
Worried About Economy	Household member(s) very worried about economy in general	0/1
Individual Health Services	Household member(s) utilized individual health services (iGel) last year	0/1
Down and Gloomy	Head of household often felt down and gloomy in the last 4 weeks	0/1
Smoker	Household member(s) is smoking	0/1
Hospital Stay	Head of household stayed at least one night in a hospital last year	0/1
Less Careful	Head of household less careful in the last 4 weeks due to emot. problems	0/1
Low Health Satisfaction	Head of household is dissatisfied with his/her health	0/1
Low Life Satisfaction	Head of household is dissatisfied with his/her life	0/1
Not Easy to Relax	Household member(s) have problems to relax after work	0/1

Table 4.2 – *Continued from previous page*

Variable	Label	Value
No Anti-Foreignism Concerns	Household member(s) not worried about hostility towards foreigners	0/1
Anti-Foreignism Concerns	Household member(s) worried about hostility towards foreigners	0/1
Peace Concerns	Head of household worried about the maintaining of peace	0/1
Strong Polit. Interest	Head of household is very interested in politics	0/1
High Risk Affinity	Head of household is willing to take risks	0/1
Sometimes Nervous	Head of household sometimes felt nervous/tense in the last 2 weeks	0/1
Refugees Mean more Risks	Household member(s) consider refugee influx risk, rather than opportunity	0/1
Refugees Negative: Economy	Household member(s) see refugees as negative influence on German Economy	0/1
Refugees Positive: Germany	Household member(s) see refugees as positive influence on Germany as place to live	0/1
Donation to Refugees	Household member(s) are willing to/did donate last year to refugees	0/1
Building, Area, and Region		
Foreigners in Area	Household in area with high share of foreigners	0/1
CDU Voters Area	Household located in area with high share of CDU voters	0/1
FDP Voters Area	Household located in area with high share of FDP voters	0/1
SPD Voters Area	Household located in area with high share of SPD voters	0/1
Supraregional Newspapers	Household located in area with high affinity towards suprareg. newsp.	0/1
High Average Age	Household located in area with high average age	0/1
High-Rise Building	Household lives in a high-rise building (9 or more stories)	0/1
High Purchasing Power	Household located in area with high purchasing power	0/1
Rhineland-Pal./Saarland	Household located in Rhineland-Palatinate or Saarland	0/1
Baden-Wuerttemberg	Household located in Baden-Wuerttemberg	0/1
Saxony	Household located in Saxony	0/1
Solar System	Dwelling is connected to solar power system	0/1
Low Status Area	Household located in area with below-average low socio-economic status	0/1
In Flat Since 2016	Household since 2016 in this dwelling	0/1
Low Share Abitur Graduates	Household located in area with low share of Abitur graduates	0/1
Refugee/Asylum Seeker Specific Variables		
Asylum Application Rejected	Household member(s): Application for asylum was rejectd, has to leave Germany	0/1
Free Elections	Household member(s) do not consider free elections part of democracy	0/1
Not First Asylum Application	Household member(s) have applied for asylum before current application	0/1
Asylum Application 2013	Household member(s) have applied for asylum until 2013	0/1
Future Education Unlikely	Head of household considers further education/training in Germany unlikely	0/1
Did Not Understand RL	Household member(s) did not understand concept of Record Linkage	0/1

Table 4.3: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2017

	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	Sample M2	Sample M3/4
Intercept	0.88***	0.03	0.96***	1.48***	1.41***	0.87***	1.09***	1.26***	0.89***	1.23***	1.21***	0.42***	1.21***	0.90***	0.55***	0.20**
Interview Characteristics																
CAPI												0.34***				
Email Known												0.24**			0.42*	
Interviewer Related HH			0.49***													
Late Interview										-0.29**						
Long Interview																-0.27*
Mode Change								-0.72**								
New Address													-0.68*			
Mother-Child-Questionnaire							-0.68**									
New Household	-1.19***		-0.85*						-0.63**		-0.92**	-0.80***	-1.00***	-0.70**		
Not Original Sample Member									-0.33***	-0.41**	-0.27**	-0.26**	-0.56***	-0.51***	-1.46***	
High Item Nonresponse P.	-0.21**								-0.68***							
Part. Unit Nonresponse										-0.31**			-0.42**			
Unrealistic Answers																-0.26**
Interview Related HH														0.47***		
Temp. Drop-Out Related HH	-0.39*													-0.55*		
Drop-Out Related HH						-0.43**	-0.85**				-1.07***	-0.42*				
Multiple Children												-0.19*				
Phone Unknown		-1.10***	-0.42***			-0.38***			-0.36*	-0.56**				-0.49**	-0.47**	-0.77***
Temporary Drop-Out	-1.29***		-1.48***		-1.84**	-1.44***		-1.46***	-0.78***	-1.42***	-1.54***	-0.82***	-1.20***	-1.18***		
Change of Interviewer	-0.37***	-0.83**				-0.42***			-0.30***		-0.47***	-0.70***	-0.50**	-0.41***	-0.34**	
SOEP Household Moved												0.34*				
Subsample M4																0.14**
Demographic Characteristics																
Age 35-44										-0.22*						
Age 65-74						0.17*		0.41*								
Parents abroad														-0.20**		
Parents in HH						-0.42*										
Family Household																0.32***
Partner not in HH														0.34**		
Female Head of HH	-0.15*															
Single Household								-0.42**						-0.29**		
Child Under 12				-1.04*												
Work, Education and Financial Situation																
Same Employer 4th Q.								-0.69***								
Receiving ALG II Benefits												0.23*				

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	Sample M2	Sample M3/4
High Number of Books						0.16*						0.20*				
Low Number of Books	-0.18*															
Low Education		0.38*														
House Owner									0.17**							
No Car in HH														0.22*		
Retirement Concerns		-0.68***														
No Retirement Concerns	0.20**															
Open-Ended Contract		-0.41*											-0.29*			
Interrupted at Work														0.26*		
Temporary Work				-1.06**												
Job Intention															0.38**	0.18**
Limited Illegal Work														0.20*		
Health, Well-Being and Personality Traits																
No Health Restrictions				-0.21*												
Angry															-0.32*	
Calm and Relaxed								0.36**								
Unhappy		0.50**														
No Health Concerns		0.51*														
Chronic Back Pain																
Overweight/High BMI				0.21*									0.21**			
Free Speech as Goal														0.20*		
Often Eat Poultry																
Never Eat Fish								-0.65**								
Worried About Economy						0.18*										
Individual Health Services															0.34*	
Down and Gloomy				-0.29*												
Smoker										0.17*					-0.17*	
Hospital Stay	0.21*											0.33**				
Less Careful						-0.33*										
Low Health Satisfaction												-0.35**				
Low Life Satisfaction															0.57**	
Not Easy to Relax						0.37**										
No Anti-Foreignism Concerns																-0.32**
Anti-Foreignism Concerns																
Peace Concerns															-0.21**	
Strong Polit. Interest														0.17*		0.22**
High Risk Affinity								0.28*								
Sometimes Nervous															0.59**	
Refugees Mean More Risks														0.21*		
Refugees Negative: Economy																-0.39*

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	Sample M2	Sample M3/4
Refugees Positive: Germany				-0.89*		0.30*										
Donation to Refugees	0.24***															
Building, Area, and Region																
Foreigners in Area			-0.36**													
CDU Voters Area								-0.42**		-0.22*		-0.29**				
FDP Voters Area													-0.40**			
SPD Voters Area															-0.23**	
Supraregional Newspapers													-0.28*			
High Average Age								-0.45***								
High-Rise Building				-0.72*											0.20*	
High Purchasing Power													0.44**			
Hesse																0.22*
Rhineland-Pal./Saarland	0.42**											-0.42**				
Baden-Wuerttemb.													-0.27*		-0.37*	
Saxony						-1.26**							-0.56***			
Solar System	-0.25*															
Low Status Area									-0.48***							
In Flat Since 2016																0.12*
Low Share Abitur Graduates																0.12*
Refugee/Asylum Seeker Specific Variables																
Asylum Application Rejected																-0.45*
Free Elections																-0.31**
Not First Asylum Application																-0.45**
Asylum Application 2013																-0.27***
Future Education Unlikely																-0.16**
Did not understand RL																-0.18*
<i>Number of Observations</i>	1,736	268	1,120	109	75	2,199	619	673	2,001	1,099	1,200	1,564	671	1,757	913	3,025
<i>Log Likelihood</i>	-551.10	-114.40	-323.97	-25.20	-15.51	-638.47	-165.19	-189.23	-624.64	-303.81	-340.60	-614.77	-269.23	-737.96	-482.84	-1586.40

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

5 Margins used in the Post-Stratification Process

In a final step, the cross-sectional weights are adjusted by a post-stratification process. The following tables provide an overview of the variables and their categories used in the post-stratification at the household level (Table 5.1) and whether they are used in a given wave and subsample (Table 5.2). Tables 5.3 and 5.4 show the same on the person level. We obtain these marginal distributions of the underlying cross-sectional population by the Microcensus provided by the Federal Statistical Office of Germany. Only in the case of marginal distributions of the IAB-BAMF-SOEP Refugee Survey, we draw on additional margins derived from the Central Register of Foreigners (AZR).

Table 5.1: Marginal Distributions - Household Level

Variables	Marginal Distributions
Federal State ²⁹ (<i>Fed. State</i>)	Berlin, Brandenburg Hamburg, Schleswig-Holstein Bremen, Lower Saxony North Rhine-Westphalia Hesse Saarland, Rhineland-Palatinate Baden-Wuerttemberg Bavaria Mecklenburg-Western Pomerania Saxony-Anhalt Thuringia Saxony
Size of Municipality (<i>Mun. Size</i>)	Less than 20,000 inhabitants 20,000-100,000 inhabitants 100,000-500,000 inhabitants More than 500,000 inhabitants
Household Size	1 — 2 — 3 — 4 — 5 or more members
Houseowner (<i>Owner</i>)	owner — tenant

²⁹Different categorisation:

Sample L1, L2 and L3: 14 units, Bremen/Hamburg and Saarland/Rhineland-Palatinate are combined

Sample J: 16 units for each Federal State

Sample M1 and M2: the last 4 units are combined in one, overall 9 categories

Table 5.1 – Continued from previous page

Variables	Marginal Distributions
Household Typology (<i>H. Type</i>)	Single household 2 adults without children 2 adults, 1 or 2 children Single parent, less than 3 children Single parent, 3 or more children Families with more than 3 children Remaining households
Migration Second Generation (<i>Migr.</i>)	No 2 nd generation migrant in household at least one 2 nd gen. migrant in household born after 1995 at least one 2 nd gen. migrant in household born 1975-1994 at least one 2 nd gen. migr. born 1975-1994 and one after 1995
Nationality (<i>Nat.</i>)	EU Country — Former Yugoslavia — Turkey — CIS countries Rest of the world — Only German nationality
Year of Immigration (<i>Imm. Year</i>)	1900-1979 — 1980-1984 — 1985-1989 — 1990-1994 1995-1999 — 2000-2004 — 2005-2009 — 2010-2013 ³⁰ — Other
Target Population AB (<i>AB</i>)	Household size and country of origin (altogether 47 combinations)
Target Population E, F (<i>E, F</i>)	West Germany, all household members German West Germany, at least one household member without Ger. nat. East Germany
Target Population G (<i>G</i>)	West Germany, household income <7,500 DM ³¹ East Germany, household income <7,500 DM West Germany, household income 7,500-10,000 DM East Germany, household income 7,500-10,000 DM West Germany, household income >10,000 DM East Germany, household income >10,000 DM

³⁰ An additional category “2010-2013” is used from 2015 on

³¹ Deutschmark (DM)

Table 5.1 – Continued from previous page

Variables	Marginal Distributions
Children Typology (Child)	Household with children aged 0-6 years Household with children aged 7-11 years Household with children aged 12-17 years Household with children aged 0-6 and children aged 7-11 Household with children aged 0-6 and children aged 12-17 Household with children aged 7-11 and children aged 12-17 Household with children aged 0-6, 7-11 and 12-17
Target Population L1 (L1)	<i>Four different variables:</i> Household with child born in 2007 (yes/no) Household with child born in 2008 (yes/no) Household with child born in 2009 (yes/no) Household with child born in the 1st quarter of 2010 (yes/no)
Target Population L2 (L2)	Family with low income (LI) Single parent household (SP) household with at least 3 children (3+) (LI) and (SP) household (LI) and (3+) household (SP) and (3+) household (LI), (SP) and (3+) household Not eligible for sample L2
Target Population L1/L2 (L1/L2)	Low income household, eligible for sample L1 Single parent household, eligible for sample L1 Household with at least 3 children, eligible for sample L1 At least 2 characteristics of sample L2 and eligible for L1 Not eligible for sample L2, but for sample L1 Eligible for sample L2, but not for sample L1 Not eligible for sample L1 and L2
Target Population L3 (L3)	Single parent household Household with at least 3 children Single parent household with at least 3 children Not eligible for sample L3
Target Population L1/L3 (L1/L3)	Single parent household, eligible for sample L1 Household with at least 3 children, eligible for sample L1 Single parent household with at least 3 children, eligible for L1 Eligible for sample L3, but not for sample L1

Table 5.1 – *Continued from previous page*

Variables	Marginal Distributions
Target Population H, J, K (<i>H, J, K</i>)	West Germany (without Berlin), all household members German West Germany, at least one household member without Ger. Nat. East Germany (incl. Berlin)
Household Size and Number of Employed Household Members ³² (<i>Empl.</i>)	Single household, not employed
	Single household, employed
	2 members, not employed
	2 members, 1 employed
	2 members, 2 employed
	3 members, not employed
	3 members, 1 employed
	3 members, 2 employed
	3 members, 3 employed
	4 or more members, not employed
	4 or more members, 1 employed
	4 or more members, 2 employed
	4 or more members, 3 employed
4 or more members, 4 or more employed	
Unemployment Benefits (<i>ALG</i>)	Household in West Germany receiving ALG II ³³
	Household in West Germany without ALG II
	Household in East Germany receiving ALG II
	Household in East Germany without ALG II
Greater Regions (<i>Reg.</i>)	North Germany — East Germany
	South Germany — West Germany

³² *Sample J*: sorted by East and West Germany

³³ Arbeitslosengeld II

Table 5.1 – Continued from previous page

Variables	Marginal Distributions
Target Population M1 (M1) ³⁴	1st Generation, 1995-2004, Turkey
	1st Generation, 1995-2004, Spain/Greece/Italy
	1st Generation, 1995-2004, Poland
	1st Generation, 1995-2004, CIS countries
	1st Generation, 1995-2004, Arabic Countries
	1st Generation, 1995-2004, Late repatriate
	1st Generation, 1995-2004, Rest of the world
	1st Generation, after 2005, Turkey, Spain, Greece
	1st Generation, after 2005, Poland
	1st Generation, after 2005, CIS countries
	1st Generation, after 2005, Rest of the world
	2nd Generation, Not Turkey
	2nd Generation, Turkey
	Target Population M2 (M2) ³³
2009-2011, Poland	
2009-2011, Romania, Bulgaria	
2009-2011, Italy, Portugal, Spain, Greece	
2009-2011, Rest of Western Europe	
2009-2011, Rest of Eastern Europe	
2009-2011, Islamic States	
2009-2011, Rest of the World	
2012-2013, Germany	
2012-2013, Poland	
2012-2013, Romania/Bulgaria	
2012-2013, Italy/Portugal/Spain/Greece	
2012-2013, Rest of Western Europe	
2012-2013, Rest of Eastern Europe	
2012-2013, Islamic States	
2012-2013, Rest of the world	
Target Population M3/4 (M3/4)	(At least one) M3/4-eligible Person moved into existing household
	Household founded by M3/4-eligible person(s)
	M3/4-eligible Person(s) living in refugee shelter
	HH not M3/4-eligible

³⁴Personal characteristics are aggregated on the household level according to the following order: 1. earliest year of immigration; 2. oldest household member; 3. female household member; 4. random household member

Table 5.1 – *Continued from previous page*

Variables	Marginal Distributions
Target Population M5 (M5)	(At least one) M5-eligible Person moved into existing household Household founded by M5-eligible person(s) M5-eligible Person(s) living in refugee shelter HH not M5-eligible
Target Population N (N)	Part of target population of Sample N Not part of target population of Sample N ³⁵
Refugee in Household (Ref.)	(At least one) Person in household came to Germany as a refugee between 2013 and 2016 No person in household that came to Germany as a refugee between 2013 and 2016

³⁵The Sample N target population consists of households in which at least one household member, on the reference date of 1 December 2011, met the following requirements: adult from 16 through 65 years of age and living in Germany.

Table 5.2: Margins - Household Level

Year (Samples)	Fed. State	Mun. Size	H. Size	Owner	H. Type	Migr.	Nat.	Imm. Year	AB	E, F	G	Child	L1	L2	L1/L2	L3	L1/L3	H, J, K	Empl.	ALG
1984 (A-B)	+ A B	+ A B	+ A B	+ A B					A B											
1985 (A-B)	+	+	+	+																
1986 (A-B)	+	+	+	+																
1987 (A-B)	+	+	+	+																
1988 (A-B)	+	+	+	+																
1989 (A-B)	+	+	+	+																
1990 (A-C)	+	+	+	+																
1991 (A-C)	+	+	+	+																
1992 (A-C)	+	+	+	+																
1993 (A-C)	+	+	+	+																
1994 (A-D)	+	+	+	+																
1995 (A-D)	+	+	+	+																
1996 (A-D)	+	+	+	+																
1997 (A-D)	+	+	+	+																
1998 (A-E)	+ * E	+ * E	+ * E	+ * E						* E										
1999 (A-E)	+	+	+	+																
2000 (A-F)	+ * F	+ * F	+ * F	+ * F						* F										
2001 (A-F)	+	+	+	+																
2002 (A-G)	+ *	+ *	+ *	+ *							* G									
2003 (A-G)	+	+	+	+																
2004 (A-G)	+	+	+	+																
2005 (A-G)	+	+	+	+																
2006 (A-H)	+ * H	+ * H	+ * H	+ * H					* H											
2007 (A-H)	+	+	+	+																
2008 (A-H)	+	+	+	+																
2009 (A-I)	+	+	+	+																
2010 (A-L2)	+ * L1 L2	+ * L1 L2	+ *	+ *	+ *	+ *						L1 L2	* L1	* L2	L1 L2					
2011 (A-L3)	+ * L3 J	+ * L3 J	+ * J	+ * J	+ * J	+ * J						L3				* L3	* L3	* J	J	J
2012 (A-K)	+ * K	+ * K	+ * K	+ * K	+ * K	+ * K												* K	K	K

Table 5.2 – Continued from previous page

Year (Samples)	Fed. State	Mun. Size	H. Size	Owner	H. Type	Migr.	Nat.	Imm. Year	Reg.	M1	M2	M3/4	M5	N	Ref.
2013 (A-M1)	+ * M1	+ * M1	+ * M1	+ *	+ *	+ *	+	+	M1	+ M1					
2014 (A-M1)	+	+	+	+	+	+	+	+							
2015 (A-M2)	+ * M2	+ * M2	+ * M2	+ *	+ *	+ *	+ *	+ *	M2		+ M2				
2016 (A-M3/4)	+ *	+ *	+ *	+ *	+ *	+ *	+ *	+ *				*			+
2017 (A-N)	+ * N	+ * N	+ * N	+ *	+ * N	+ * N	+ * N	+ * N					*	*	+

Note: (+) margins for standard weights; (*) margins for standard weights without the new samples; (*sample letter*) margins for standalone weights of a new sample

Table 5.3: Marginal Distributions - Person Level

Variables	Distributions
Age and Gender	0-4 male — 0-4 female — 5-9 male — 5-9 female 10-14 male — 10-14 female — 15-19 male — 15-19 female 20-24 male — 20-24 female — 25-29 male — 25-29 female 30-34 male — 30-34 female — 35-39 male — 35-39 female 40-44 male — 40-44 female — 45-49 male — 45-49 female 50-54 male — 50-54 female — 55-59 male — 55-59 female 60-64 male — 60-64 female — 65-69 male — 65-69 female 70+ male — 70+ female
Household Typology (<i>H. Type</i>)	1 adult and 0 children — 2 adults and 0 children 3 adults and 0 children — 4 or more adults and 0 children 1 adult and 1 or more children — 2 adults and 1 child 2 adults and 2 children — 2 adults and 3 or more children 3 adults and 1 or more children 4 or more adults and 1 or more children
German Nationality (<i>German</i>)	German nationality — Other nationality
Target Population G (<i>G</i>)	West Germany, household income <7,500 DM ³⁶ East Germany, household income <7,500 DM West Germany, household income 7,500-10,000 DM East Germany, household income 7,500-10,000 DM West Germany, household income >10,000 DM East Germany, household income >10,000 DM
Migration Second Generation (<i>Migrant 2nd Gen.</i>)	Indirect migration, born after 1995 Indirect migration, German nat., born 1975/1994 Indirect migration, other nat., born 1975/1994 Indirect migration, other nat. born before 1964 until 1974 Direct or no migration, or indirect migration, but German nationality and born before 1975
Age ³⁷	0-4 — 5-9 — 10-14 — 15-19 — 20-24 — 25-29 — 30-34 35-39 — 40-44 — 45-49 — 50-54 — 55-59 — 60-64 — 65+
Gender	Male — Female

³⁶Deutschmark (DM)

³⁷Different categorisation in:

Sample L1: 0, 1, 2, 3, 4-7, 8-12, 13-18, 19-26, 27-31, 32-36, 37-41, 42-46, 47+

Sample L2: 0-3, 4-7, 8-12, 13-18, 19-26, 27-31, 32-36, 37-41, 42-46, 47-51, 52-56, 57+

Sample L3: 0-3, 4-6, 7-11, 12-17, 18-25, 26-30, 31-35, 36-40, 41-45, 46-50, 51-55, 56+

Sample M1: For respondents younger than 19 years old: only one category (0-19).

Table 5.3 – *Continued from previous page*

Variables	Distributions
Target Population L1 (L1)	<p><i>Four different variables:</i></p> <p>Household with child born in 2007 (yes/no)</p> <p>Household with child born in 2008 (yes/no)</p> <p>Household with child born in 2009 (yes/no)</p> <p>Household with child born in the 1st quarter of 2010 (yes/no)</p>
Target Population L2 (L2)	<p>Family with low income (LI)</p> <p>Single parent household (SP)</p> <p>Household with at least 3 children (3+)</p> <p>(LI) and (SP) household</p> <p>(LI) and (3+) household</p> <p>(SP) and (3+) household</p> <p>(LI), (SP) and (3+) household</p>
Target Population L3 (L3)	<p>Single parent household (SP)</p> <p>Household with at least 3 children (3+)</p> <p>(SP) and (3+) household</p>
Foreign Nationality (Nation.)	<p>EU Country — Former Yugoslavia — CIS countries — Turkey</p> <p>Rest of the world — Only German nationality</p>
Year of Immigration (Imm. Year)	<p>1900-1979 — 1980-1984 — 1985-1989 — 1990-1994</p> <p>1995-1999 — 2000-2004 — 2005-2009 — 2010-2013³⁸ — Other</p>
Type of Migration Background (Migrant)	<p>Immigration before 1995</p> <p>Immigration between 1995 and 2004</p> <p>Immigration since 2005</p> <p>Migration background (indirect)</p> <p>No migration background</p> <p>Not eligible for sample M1</p>

³⁸An additional category “2010-2013” is used from 2015 on.

Table 5.3 – *Continued from previous page*

Variables	Distributions
Target Population M1 (M1)	1st generation, earlier than 1995, Turkey, m/f ³⁹
	1st generation, earlier than 1995, Spain/Greece/Italy, m/f
	1st generation, earlier than 1995, Late repatriate, m/f
	1st generation, earlier than 1995, Rest of the world, m/f
	1st generation, 1995-2004, Turkey, m/f
	1st generation, 1995-2004, Spain/Greece/Italy, m/f
	1st generation, 1995-2004, Poland, m/f
	1st generation, 1995-2004, CIS countries, m/f
	1st generation, 1995-2004, Arabic countries, m/f
	1st generation, 1995-2004, Late repatriate, m/f
	1st generation, 1995-2004, Rest of the world, m/f
	1st generation, after 2005, Spain/Greece/Italy, m/f
	1st generation, after 2005, Poland, m/f
	1st generation, after 2005, CIS countries, m/f
	1st generation, after 2005, Rest of the world, m/f
	2nd generation, Not Turkey, m/f
	2nd generation, Turkey, m/f
	German, m/f
	Not eligible for sample M1
	Target Population M2 (M2)
2009-2011, Germany, m/f	
2009-2011, Poland, m/f	
2009-2011, Romania/Bulgaria, m/f	
2009-2011, Italy/Portugal/Spain/Greece, m/f	
2009-2011, Rest of Western Europe, m/f	
2009-2011, Rest of Eastern Europe, m/f	
2009-2011, Islamic States, m/f	
2009-2011, Rest of the world, m/f	
2012-2013, Germany, m/f	
2012-2013, Poland, m/f	
2012-2013, Romania/Bulgaria, m/f	
2012-2013, Italy/Portugal/Spain/Greece, m/f	
2012-2013, Rest of Western Europe, m/f	
2012-2013, Rest of Eastern Europe, m/f	
2012-2013, Islamic States, m/f	
2012-2013, Rest of the world, m/f	

³⁹Each category distinguishes between male (m) or female (f) gender of the respondent.

Table 5.3 – *Continued from previous page*

Variables	Distributions
Federal State - Refugee Samples ⁴⁰ (<i>Ref. Fed. State</i>)	Berlin, Brandenburg Hamburg, Schleswig-Holstein Bremen, Lower Saxony North Rhine-Westphalia Hesse Saarland, Rhineland-Palatinate Baden-Wuerttemberg Bavaria Mecklenburg-Western Pomerania Saxony-Anhalt Thuringia Saxony
Registered Date of Arrival - Refugee Samples (<i>Ref. Reg.</i>)	Arrival until/including January 2016 Arrival from February 2016 on
Date of Arrival in Germany ⁴¹ - Refugee Samples - By Year and Quarter (<i>Ref. Arrival</i>)	2013 Q1 — 2013 Q2 — 2013 Q3 — 2013 Q4 2014 Q1 — 2014 Q2 — 2014 Q3 — 2014 Q4 2015 Q1 — 2015 Q2 — 2015 Q3 — 2015 Q4 2016 Q1 — 2016 Q2 — 2016 Q3 — 2016 Q4
Country of Origin - Refugee Samples (<i>Ref. Origin</i>)	Syria Afghanistan Iraq Albania, Serbia, Kosovo Eritrea, Somalia Iran, Pakistan Other
Age - Refugee Samples (<i>Ref. Age</i>)	0-4 — 5-9 — 10-14 — 15-17 — 18-20 — 21-24 — 25-29 — 30-34 35-39 — 40-44 — 45-49 — 50-54 — 55-59 — 60+
Gender - Refugee Samples (<i>Ref. Gender</i>)	Male — Female

⁴⁰The variables marked here with "*Refugee Samples*" refer to Samples M3/4 in 2016 and Samples M3/4 and M5 in 2017 respectively.

⁴¹The date of arrival in this variable is based on self-reported information. This information may differ from the officially *registered* date of arrival recorded in the corresponding variable above.

Table 5.3 – *Continued from previous page*

Variables	Distributions
Refugee (<i>Ref.</i>)	Came to Germany as a refugee between 2013 and 2016 Did not come to Germany as a refugee between 2013 and 2016

Table 5.4: Margins - Person Level

Year (Samples)	Age and Gender	H. Type	German	G	Migrant 2 nd Gen.	Age	Gender	L1	L2	L3	Nation	Imm. Year	Migrant	M1	M2	Ref. Fed. State	Ref. Reg.	Ref. Arrival	Ref. Origin	Ref. Age	Ref. Gender	Ref.
1984 (A-B)	+	+	+																			
1985 (A-B)	+	+	+																			
1986 (A-B)	+	+	+																			
1987 (A-B)	+	+	+																			
1988 (A-B)	+	+	+																			
1989 (A-B)	+	+	+																			
1990 (A-B)	+	+	+																			
1991 (A-B)	+	+	+																			
1992 (A-B)	+	+	+																			
1993 (A-B)	+	+	+																			
1994 (A-B)	+	+	+																			
1995 (A-B)	+	+	+																			
1996 (A-B)	+	+	+																			
1997 (A-B)	+	+	+																			
1998 (A-E)	+ * E	+ * E	+ * E																			
1999 (A-E)	+	+	+																			
2000 (A-F)	+ * F	+ * F	+ * F																			
2001 (A-F)	+	+	+																			
2002 (A-H)	+ *	+ *	+ *	* G																		
2003 (A-H)	+	+	+																			
2004 (A-H)	+	+	+																			
2005 (A-H)	+	+	+																			
2006 (A-H)	+ * H	+ * H	+ * H																			
2007 (A-H)	+	+	+																			
2008 (A-H)	+	+	+																			
2009 (A-I)	+	+	+																			
2010 (A-L2)	+ *	+ *	+ *		+ *	L1 L2	L1 L2	L1	L2													
2011 (A-L3)	+ * J	+ * J	+ * J		+ * J	L3	L3			L3												
2012 (A-K)	+ * K	+ * K	+ * K		+ * K																	

Note. (+) margins for standard weights; (*) margins for standard weights without the new samples; (sample letter) margins for standalone weights of a new sample

Table 5.4 – *Continued from previous page*

Year (Samples)	Age and Gender	H. Type	German	G	Migrant 2 nd Gen.	Age	Gender	L1	L2	L3	Nation	Imm. Year	Migrant	M1	M2	Ref. Fed. State	Ref. Reg.	Ref. Arrival	Ref. Origin	Ref. Age	Ref. Gender	Ref.
2013 (A-M1)	+ *	+ *	+ *		+ *	M1					+ *	+ *	M1	* M1								
2014 (A-M1)	+	+	+		+						+	+										
2015 (A-M2)	+ *	+ *	+ *		+ *	M2					+ *	+ *			M2							
2016 (A-M3/4)	+ *	+ * M3/4	+ *		+ *						+ *	+ *				+ M3/4		+ M3/4	M3/4	+ M3/4	+ M3/4	+
2017 (A-N)	+ * N	+ * M5	+ * N		+ * N						+ * N	+ * N				+ M5	+ M5	+ M5	+ M5	+ M5	+ M5	+

Note. (+) margins for standard weights; (*) margins for standard weights without the new samples;
(*sample letter*) margins for standalone weights of a new sample

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

Based on the regression models of successful vs. unsuccessful re-contacts 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 2017 based on characteristics measured in 2016, variable `BHHBLEIB`, lends itself as a longitudinal weighting variable which itself corrects for selective attrition between waves 2016 and 2017. Tables 6.1, Table 6.2, Table 6.3, Table 6.4 and Table 6.5 report some subsample specific summary statistics of the longitudinal weights in each wave.

The product of the cross-sectional weight in 2016, variable `BGHHRF`, and the longitudinal weight in 2017, variable `BHHBLEIB`, provide the raw data for the cross-sectional weight in 2017. 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 2017.

Tables 6.6 and 6.7 report subsample specific summary statistics of the derived cross-sectional weighting variable `BHHHRF` and in comparison all previous cross-sectional weights `AHHRF` through `BGHHRF`.

Table 6.1: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples A through D (Percentiles of \$HBLEIB up to Wave 34).

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
2016	1.03	1.08	1.24	1,629	1.01	1.10	1.86	228	1.01	1.07	1.30	1,073	1.01	1.07	1.43	103
2017	1.02	1.09	1.24	1,528	1.02	1.17	1.79	201	1.02	1.08	1.22	997	1.02	1.02	1.28	99

Table 6.2: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples E through G (Percentiles of \$HBLEIB up to Wave 34).

Year	Sample E				Sample F				Sample G			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
1998												
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
2016	1.06	1.06	1.38	68	1.03	1.08	1.24	2,094	1.02	1.02	1.26	590
2017	1.02	1.02	1.45	67	1.03	1.10	1.25	1,968	1.02	1.06	1.22	561

Table 6.3: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples H, J and K (Percentiles of \$HBLEIB up to Wave 34).

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
2016	1.01	1.04	1.29	639	1.02	1.06	1.20	1,883	1.02	1.05	1.27	1,046
2017	1.01	1.05	1.35	594	1.06	1.10	1.22	1,776	1.03	1.07	1.20	987

Table 6.4: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples L1, L2 and L3 (Percentiles of \$HBLEIB up to Wave 34).

Year	Sample L1				Sample L2				Sample L3			
	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	1.02	1.11	1.47	1,247	1.10	1.26	1.67	1,416	1.10	1.25	1.76	593
2015	1.01	1.06	1.36	1,184	1.04	1.15	1.91	1,379	1.03	1.12	1.74	582
2016	1.02	1.08	1.25	1,122	1.05	1.16	1.97	1,265	1.03	1.15	1.66	533
2017	1.02	1.06	1.24	1,055	1.03	1.12	1.64	1,247	1.03	1.10	1.93	516

Table 6.5: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples M1 and M2 (Percentiles of \$HBLEIB up to Wave 34).

Year	Sample M1				Sample M2				Sample M3/4			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
2014	1.08	1.28	1.81	2,012								
2015	1.08	1.27	1.89	1,667								
2016	1.07	1.21	1.90	1,493	1.24	1.50	2.23	660				
2017	1.03	1.16	1.61	1,350	1.10	1.36	3.02	559	1.10	1.31	2.06	2,178

Table 6.6: Summary Statistics of the Derived Cross-Sectional Weights at the Household Level (Percentiles of \$HHRF up to Wave 34).

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	483	683	3,899	5,081	6,429	8,471	10,033	5,322
1986	539	752	3,597	5,301	6,839	9,280	11,117	5,090
1987	546	790	3,546	5,381	7,044	9,576	11,455	5,026
1988	532	804	3,561	5,638	7,542	10,355	12,539	4,814
1989	549	819	3,602	5,840	7,879	10,810	13,274	4,690
1990	699	1,073	2,217	4,601	7,042	9,897	12,393	6,819
1991	680	1,043	2,326	4,692	7,144	10,288	12,873	6,699
1992	668	1,028	2,334	4,662	7,139	10,530	13,649	6,665
1993	688	1,056	2,400	4,669	7,259	10,750	13,978	6,637
1994	708	1,102	2,401	4,674	7,284	11,206	14,713	6,559
1995	696	1,113	2,387	4,362	6,976	11,089	14,840	6,768
1996	732	1,162	2,385	4,359	7,010	11,381	15,309	6,699
1997	736	1,206	2,400	4,318	7,049	11,850	15,874	6,621
1998	979	1,354	2,331	3,975	6,225	9,891	13,119	7,492
1999	968	1,320	2,311	3,987	6,502	10,885	14,345	7,220
2000	799	1,101	1,758	2,525	3,569	5,083	6,512	13,082
2001	752	1,029	1,755	2,752	4,145	6,095	7,831	11,796
2002	506	658	1,221	2,555	4,192	6,513	8,256	12,320
2003	502	673	1,237	2,562	4,327	6,827	9,084	11,909
2004	491	669	1,214	2,535	4,420	7,264	9,830	11,644
2005	490	678	1,235	2,544	4,519	7,586	10,872	11,294
2006	458	650	1,268	2,391	4,139	6,878	9,762	12,361
2007	457	652	1,251	2,467	4,468	7,587	10,675	11,552
2008	460	655	1,277	2,556	4,752	8,243	11,572	10,921
2009	473	668	1,302	2,630	5,032	9,091	12,358	10,270
2010	220	360	667	1,429	3,645	7,366	11,107	13,888
2011	215	326	610	1,503	3,098	5,605	7,821	16,703
2012	218	329	636	1,635	3,163	5,704	7,621	16,397
2013	174	270	519	1,294	2,938	5,391	7,430	17,992
2014	200	311	618	1,518	3,367	6,260	8,662	15,946
2015	186	299	615	1,475	3,388	6,539	9,126	15,908
2016	38	76	338	1,153	3,112	6,136	9,143	17,715
2017	37	70	316	1,142	2,801	5,438	8,169	19,628

Table 6.7: Summary Statistics of the Derived Cross-Sectional Weights at the Person Level (Percentiles of \$PHRF up to Wave 34).

Year	p5	p10	p25	p50	p75	p90	p95	N
1984	397	553	1,172	4,365	5,227	6,053	6,800	16,173
1985	453	634	1,421	4,621	5,718	6,893	8,065	14,508
1986	488	678	1,532	4,679	6,024	7,591	9,011	13,804
1987	509	719	1,598	4,727	6,223	7,893	9,407	13,563
1988	487	685	1,622	4,897	6,565	8,501	10,212	12,872
1989	528	755	1,747	5,018	6,892	8,969	10,691	12,443
1990	681	1,023	1,900	3,440	6,146	8,286	10,234	18,254
1991	731	1,072	1,915	3,707	6,191	8,478	10,605	17,844
1992	780	1,139	1,998	3,739	6,301	8,726	11,117	17,429
1993	843	1,233	2,087	3,835	6,385	9,001	11,384	17,072
1994	872	1,283	2,107	3,843	6,417	9,273	12,063	16,715
1995	766	1,144	2,017	3,610	6,100	9,071	12,157	17,345
1996	797	1,190	2,023	3,642	6,144	9,415	12,771	16,944
1997	838	1,215	2,063	3,668	6,250	9,697	13,351	16,583
1998	913	1,267	2,041	3,512	5,595	8,508	11,403	18,249
1999	902	1,240	2,015	3,494	5,798	9,242	12,667	17,501
2000	723	973	1,565	2,314	3,214	4,561	5,852	30,784
2001	690	935	1,535	2,452	3,646	5,409	6,931	27,956
2002	446	612	1,064	2,193	3,713	5,817	7,630	29,101
2003	440	620	1,086	2,204	3,813	6,121	8,220	27,867
2004	436	616	1,086	2,186	3,893	6,512	8,849	26,918
2005	440	630	1,117	2,235	4,013	6,876	9,605	25,638
2006	414	597	1,116	2,176	3,680	6,293	8,819	27,442
2007	415	599	1,123	2,230	3,924	6,952	10,081	25,505
2008	425	611	1,154	2,298	4,129	7,642	11,245	23,792
2009	440	627	1,182	2,387	4,361	8,359	12,530	22,096
2010	175	278	535	1,033	2,519	5,410	8,561	35,945
2011	165	252	456	986	2,388	4,426	6,607	42,031
2012	167	250	470	1,099	2,544	4,508	6,719	40,351
2013	139	212	409	893	2,236	4,300	6,356	44,633
2014	156	243	480	1,081	2,590	4,935	7,328	38,839
2015	143	229	471	1,086	2,585	5,101	7,609	38,224
2016	21	37	185	792	2,297	4,843	7,339	44,042
2017	23	37	160	783	2,158	4,394	6,613	48,249

References

- AAPOR (2016). *Standard Definitions. Final Dispositions of Case Codes and Outcome Rates for Surveys*. Tech. rep. American Association of Public Opinion.
- Eisnecker, P. S. and M. Kroh (2017). “The Informed Consent to Record Linkage in Panel Studies: Optimal Starting Wave, Consent Refusals, and Subsequent Panel Attrition”. In: *Public Opinion Quarterly* 81.1, 131–143.
- Eisnecker, P. S., K. Erhardt, M. Kroh, and P. Trübswetter (2017). *The Request for Record Linkage in the IAB-SOEP Migration Sample*. SOEP Survey Papers 291, DIW/SOEP, Berlin 2017.
- Infratest Sozialforschung (2011a). *SOEP 1984 – Methodenbericht zum Befragungsjahr 1984 (Welle 1) des Sozio-oekonomischen Panels*. SOEP Survey Papers 1, DIW/SOEP, Berlin 2011.
- (2011b). *SOEP 1990/91 – Methodenbericht Ostdeutschland zu den Befragungsjahren 1990-1991 (Welle 1/2 – Ost) des Sozio-oekonomischen Panels*. SOEP Survey Papers 14, DIW/SOEP, Berlin 2011.
- (2011c). *SOEP 1994 – Methodenbericht Zuwanderer-Befragung (Teilstichprobe D1) zum Befragungsjahr 1994 (Welle 11) des Sozio-oekonomischen Panels*. SOEP Survey Papers 26, DIW/SOEP, Berlin 2011.
- (2011d). *SOEP 1995 – Methodenbericht Zuwanderer-Befragung II (Zweitbefragung D1, Erstbefragung D2) zum Befragungsjahr 1995 (Welle 12) des Sozio-oekonomischen Panels*. SOEP Survey Papers 28, DIW/SOEP, Berlin 2011.
- (2011e). *SOEP 1998 – Methodenbericht Erstbefragung der Stichprobe E zum Befragungsjahr 1998 (Welle 15) des Sozio-oekonomischen Panels*. SOEP Survey Papers 33, DIW/SOEP, Berlin 2011.
- (2011f). *SOEP 2000 – Methodenbericht erste Welle der SOEP Stichprobe F zum Befragungsjahr 2000 (Welle 17) des Sozio-oekonomischen Panels*. SOEP Survey Papers 37, DIW/SOEP, Berlin 2011.
- (2011g). *SOEP 2002 – Methodenbericht Sondererhebung Hocheinkommensstichprobe zum Befragungsjahr 2002 (Welle 19) des Sozio-oekonomischen Panels*. SOEP Survey Papers 44, DIW/SOEP, Berlin 2011.
- (2011h). *SOEP 2003 – Methodenbericht zweite Welle der Sondererhebung Hocheinkommensstichprobe zum Befragungsjahr 2003 (Welle 20) des Sozio-oekonomischen Panels*. SOEP Survey Papers 47, DIW/SOEP, Berlin 2011.
- Jacobsen, J., M. Kroh, S. Kühne, J. A. Scheible, R. Siegers, and M. Siegert (2019). *Supplementary of the IAB-BAMF-SOEP Survey of Refugees in Germany (M5) 2017*. SOEP Survey Papers 605, DIW/SOEP, Berlin 2019.
- Kara, S., S. Zimmermann, and SOEP-Group (2018). *SOEPcompanion (v34), Release 2018, v.1* SOEP Survey Papers 588, DIW/SOEP, Berlin 2018.
- Kühne, S. and M. Kroh (2017). *The 2015 IAB-SOEP Migration Study M2: Sampling Design, Nonresponse, and Weighting Adjustment*. SOEP Survey Papers 473, DIW/SOEP, Berlin 2017.
- Kroh, M., K. Käppner, and S. Kühne (2014). *Sampling, Nonresponse, and Weighting in the 2011 and 2012 Refreshment Samples J and K of the Socio-Economic Panel*. SOEP Survey Papers 260, DIW/SOEP, Berlin 2014.

- Kroh, M., R. Siegers, and S. Kühne (2015a). *Gewichtung und Integration von Auffrischungsstichproben 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.
- Kroh, M., S. Kühne, J. Goebel, and F. Preu (2015b). *The 2013 IAB-SOEP Migration Sample (M1): Sampling Design and Weighting Adjustment*. SOEP Survey Papers 271, DIW/SOEP, Berlin 2015.
- Kroh, M., H. Brücker, S. Kühne, E. Liebau, J. Schupp, M. Siegert, and P. Trübswetter (2016). *Das Studiendesign der IAB-BAMF-SOEP-Befragung von Geflüchteten*. SOEP Survey Papers 365, DIW/SOEP, Berlin 2016.
- Kroh, M., S. Kühne, J. Jacobsen, M. Siegert, and R. Siegers (2017). *Sampling, Nonresponse, and Integrated Weighting of the 2016 IAB-BAMF-SOEP Survey of Refugees (M3/M4) – revised version*. SOEP Survey Papers 477, DIW/SOEP, Berlin 2017.
- Kroh, M., V. Belcheva, and R. Siegers (2018). *Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2016)*. SOEP Survey Papers 480, DIW/SOEP, Berlin 2018.
- Martin, S., A. Zabal, and B. Rammstedt (2018). “PIAAC-L data collection 2016: technical report.” In: *GESIS Papers 2018*|05.
- OECD (2016). *Technical Report of the Survey of Adult Skills (PIAAC), 2nd Edition*. Tech. rep. OECD - Organisation for Economic Co-operation and Development.
- Pforr, K., M. Blohm, A. G. Blom, B. Erdel, B. Felderer, M. Fräßdorf, K. Hajek, S. Helmschrott, C. Kleinert, A. Koch, U. Krieger, M. Kroh, S. Martin, D. Saßenroth, C. Schmiedeberg, E.-M. Trüdinger, and B. Rammstedt (2015). “Are Incentive Effects on Response Rates and Nonresponse Bias in Large-scale, Face-to-face Surveys Generalizable to Germany? Evidence from Ten Experiments”. In: *Public Opinion Quarterly* 79.3, 740–768.
- Projektgruppe Das Sozio-oekonomische Panel (DIW) (1998). *Funktion und Design einer Ergänzungsstichprobe für das Sozio-oekonomische Panel (SOEP)*. DIW Discussion Papers 163, Berlin 1998.
- Rendtel, U. (1995). *Lebenslagen im Wandel: Panellausfälle und Panelrepräsentativität*. Vol. 8. Campus Verlag.
- Rendtel, U., M. Pannenberg, and S. Daschke (1997). “Die Gewichtung der Zuwanderer-Stichprobe des Sozio-oekonomischen Panels (SOEP)”. In: *Vierteljahrshefte zur Wirtschaftsforschung* 66.2, pp. 271–286.
- Schonlau, M., N. Watson, and M. Kroh (2011). “Household survey panels: how much do following rules affect sample size?” In: *Survey Research Methods* 5.2, pp. 53–61.
- 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.
- Schräpler, J.-P., J. Schupp, and G. G. Wagner (2006). *Changing From PAPI to CAPI – A Longitudinal Study of Mode Effects Based on an Experimental Design*. DIW Discussion Papers 593, Berlin 2006.
- Schröder, M., D. Saßenroth, J. Körtner, M. Kroh, and J. Schupp (2013a). *Experimental Evidence of the Effect of Monetary Incentives on Cross-Sectional and Longitudinal Response: Experiences from the Socio-Economic Panel (SOEP)*. SOEPpapers 603, DIW/SOEP, Berlin 2013.

- Schröder, M., R. Siegers, and C. K. Spieß (2013b). “Familien in Deutschland – FiD”. In: *Schmollers Jahrbuch* 133.4, pp. 595–606.
- 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.
- TNS Infratest Sozialforschung (2011a). *”Familien in Deutschland” (FiD) 2010 Methodenbericht: Anlage und Ergebnisse der FiD-Stichproben*. München 2011.
- (2011b). *”Familien in Deutschland” (FiD) 2011 Methodenbericht: Anlage und Ergebnisse der FiD-Stichproben*. München 2011.
- (2011c). *SOEP 2006 – Methodenbericht Erstbefragung der Ergänzungsstichprobe H zum Befragungsjahr 2006 (Welle 23) des Sozio-oekonomischen Panels*. SOEP Survey Papers 57, DIW/SOEP, Berlin 2011.
- (2012a). *SOEP 2009 – Methodenbericht Innovationssample zum Befragungsjahr 2009 (Welle 26) des Sozio-oekonomischen Panels (Erstbefragung Stichprobe I)*. SOEP Survey Papers 73, DIW/SOEP, Berlin 2012.
- (2012b). *SOEP 2011 – Methodenbericht zum Befragungsjahr 2011 (Welle 28) des Sozio-oekonomischen Panels*. SOEP Survey Papers 108, DIW/SOEP, Berlin 2012.
- (2013). *SOEP 2012 - Methodenbericht zum Befragungsjahr 2012 (Welle 29) des Sozio-oekonomischen Panels*. SOEP Survey Papers 144, DIW/SOEP, Berlin 2013.
- (2014). *Methodenbericht zum IAB-SOEP-Migrationssample 2013*. SOEP Survey Papers 217, DIW/SOEP, Berlin 2014.
- Zabal, A., S. Martin, N. Massing, D. Ackermann, S. Helmschrott, I. Barkow, and B. Rammstedt (2014). *PIAAC Germany 2012. Technical Report*. Tech. rep. GESIS - Leibniz Institute for Social Sciences.
- Zabal, A., S. Martin, and B. Rammstedt (2016). “PIAAC-L data collection 2014: technical report; follow-up to PIAAC Germany 2012.” In: *GESIS Papers* 2016|17.
- (2017). “PIAAC-L data collection 2015: technical report.” In: *GESIS Papers* 2017|29.