

## SOEP Survey Papers

Series C – Data Documentation

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

2018

# SOEP-Core – Documentation of Sample Sizes and Panel Attrition (1984 until 2016)

Martin Kroh, Simon Kühne, Rainer Siegers, Veronika Belcheva

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:

Martin Kroh, Simon Kühne, Rainer Siegers, Veronika Belcheva. 2018. SOEP-Core – Documentation of Sample Sizes and Panel Attrition (1984 until 2016). SOEP Survey Papers 480: Series C. Berlin: DIW/SOEP



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

© 2018 by SOEP

ISSN: 2193-5580 (online)

DIW Berlin

German Socio-Economic Panel (SOEP)

Mohrenstr. 58

10117 Berlin

Germany

soeppapers@diw.de

# **SOEP-Core – Documentation of Sample Sizes and Panel Attrition (1984 until 2016)**

**Martin Kroh, Simon Kühne, Rainer Siegers, Veronika Belcheva**

# Contents

<b>1</b>	<b>Introduction</b>	<b>6</b>
<b>2</b>	<b>Sampling of SOEP Subsamples A to M3/4</b>	<b>8</b>
2.1	Sample A (1984)	8
2.2	Sample B (1984)	9
2.3	Sample C (1990)	10
2.4	Sample D (1994/95)	11
2.5	Sample E (1998)	12
2.6	Sample F (2000)	13
2.7	Sample G (2002)	14
2.8	Sample H (2006)	15
2.9	Sample I (2009)	16
2.10	Sample J (2011)	17
2.11	Sample K (2012)	18
2.12	Sample L1 (FiD) (2010)	19
2.13	Sample L2 (FiD) (2010)	20
2.14	Sample L3 (FiD) (2011)	21
2.15	Sample M1 (2013)	22
2.16	Sample M2 (2015)	23
2.17	Sample M3/4 (2016)	24
<b>3</b>	<b>Developments in Sample Size</b>	<b>25</b>
3.1	Development of the Number of Successful Interviews by Cross-Section	25
3.2	Continuance and Exit: The First Wave Gross Samples and their Participation Behavior	39
3.3	New Entrants through birth or move into SOEP Households and their Participation Behavior	41
3.4	Original Households and Split-Offs	43
3.5	The Risk of Survey-Related Panel Attrition	45
<b>4</b>	<b>Panel Attrition Due to Unsuccessful Follow-Ups</b>	<b>50</b>
4.1	The Frequency of Successful Follow-Ups	50
4.2	Predicting the Probability of Successful vs. Unsuccessful Follow-Ups in the Year 2016	52
<b>5</b>	<b>Panel Attrition Due to Refusals</b>	<b>56</b>
5.1	The Frequency of Participation	56
5.2	Predicting the Probability of Re-Interviewing vs. Refusal in the Year 2016	58
<b>6</b>	<b>Margins used in the Post-Stratification Process</b>	<b>66</b>
<b>7</b>	<b>Summary Statistics of the Derived Longitudinal and Cross-Sectional Weights</b>	<b>80</b>

## List of Tables

4.1	The Frequency of Households to be Re-Contacted and the Percentage of Successful Follow-Ups, Subsamples A to M3/4 by Year. . . . .	51
4.2	Definition of the Regressors of the Logit Model of Unsuccessful Follow-Ups . . . . .	53
4.3	Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2016 . . . . .	54
5.1	The Frequency of Re-Contacted Households and the Percentage of Participation, Subsamples A to M3/4 by Year. . . . .	57
5.2	Definition of the Regressors of the Logit Model of Refusal . . . . .	59
5.3	Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2016 . . . . .	62
6.1	Marginal Distributions - Household Level . . . . .	66
6.2	Margins - Household Level . . . . .	71
6.3	Marginal Distributions - Person Level . . . . .	73
6.4	Margins - Person Level . . . . .	78
7.1	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples A through D (Percentiles of \$HBLEIB up to Wave BG). . . . .	81
7.2	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples E through G (Percentiles of \$HBLEIB up to Wave BG). . . . .	82
7.3	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples H, J and K (Percentiles of \$HBLEIB up to Wave BG). . . . .	83
7.4	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples L1, L2 and L3 (Percentiles of \$HBLEIB up to Wave BG). . . . .	83
7.5	Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples M1 and M2 (Percentiles of \$HBLEIB up to Wave BG). . . . .	83
7.6	Summary Statistics of the Derived Cross-Sectional Weights at the Household Level (Percentiles of \$HHRF up to Wave 33). . . . .	84
7.7	Summary Statistics of the Derived Cross-Sectional Weights at the Person Level (Percentiles of \$PHRF up to Wave 33). . . . .	85

## List of Figures

1	The Number of Successful Interviews with Persons by Subsamples A through M3/4, Waves 1 to 33. . . . .	25
2	Comparison of Successful Interviews with Persons and Households (Subsamples A and B), Waves 1 to 33 . . . . .	26
3	Comparison of Successful Interviews with Persons and Households (Subsamples C), Waves 1 to 27 . . . . .	27
4	Comparison of Successful Interviews with Persons and Households (Subsamples D), Waves 1 to 22 . . . . .	28
5	Comparison of Successful Interviews with Persons and Households (Subsample E), Waves 1 to 19 . . . . .	29
6	Comparison of Successful Interviews with Persons and Households (Subsample F), Waves 1 to 17 . . . . .	30
7	Comparison of Successful Interviews with Persons and Households (Subsample G), Waves 1 to 15 . . . . .	31
8	Comparison of Successful Interviews with Persons and Households (Subsample H), Waves 1 to 11 . . . . .	32
9	Comparison of Successful Interviews with Persons and Households (Subsample J), Waves 1 to 6 . . . . .	33
10	Comparison of Successful Interviews with Persons and Households (Subsample K), Waves 1 to 5 . . . . .	34
11	Comparison of Successful Interviews with Persons and Households (Subsample L1), Waves 1 to 7 . . . . .	35
12	Comparison of Successful Interviews with Persons and Households (Subsample L2), Waves 1 to 7 . . . . .	36
13	Comparison of Successful Interviews with Persons and Households (Subsample L3), Waves 1 to 6 . . . . .	37
14	Comparison of Successful Interviews with Persons and Households (Subsample M1), Waves 1 to 4 . . . . .	38
15	First-Wave Persons and their Participation Behavior. Development up to 2016 .	39
16	Entrants and their Participation Behavior. Development up to 2016 . . . . .	41
17	Proportion of First-Wave and New Households. Development up to 2016 . . . .	43
18	Successful Re-Interviewing of First-Wave Respondents by Subsamples A, B, C. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad . . . . .	45
19	Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad . . . . .	46
20	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 . . . . .	46
21	Successful Re-Interviewing of First-Wave Respondents by Subsamples L1, L2, L3 and M1. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad . . . . .	47

22	Successful Re-Interviewing of All First-Wave Respondents by Age Categories. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad . . . . .	47
23	Successful Re-Interviewing of All First-Wave Respondents by Occupation. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	48
24	Successful Re-Interviewing of All First-Wave Respondents by Income Quintiles. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad . . . . .	48
25	Successful Re-Interviewing of All First-Wave Respondents by Education. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad	49

# 1 Introduction

This data documentation is meant to provide SOEP users with a general overview of the longitudinal development of the survey over the past 33 years and the derivation of weights that compensate for disproportional sampling probabilities, selective non-response 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, and the 2016 IAB-BAMF-SOEP Refugee Sample M3/4. 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, Haisken-DeNew and Frick (2017)). 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 2016 based on the characteristics of households measured in 2015. 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. (2017a) for wave bf).

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 2016: BGHBLEIB and BGPBLEIB. Based on the



inverse of the probability of observing households and persons in 2015, the staying probability in 2016, and additional post-stratification to meet benchmarks of known margins of the underlying population in 2016, we derive the cross-sectional weights  $BGHHRF$  and  $BGPHRF$ .

Section 6 illustrates which margins are used during the post-stratification process in the different waves. Especially samples L1-L3 and M1-M3/4 that cover specific sub-populations have required to modify the selection and coding of margins over time.

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.

## 2 Sampling of SOEP Subsamples A to M3/4

### 2.1 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

---

<b>Sampling Design</b>	Multistage stratified sampling procedure based on the ADM-Design <sup>1</sup>		
<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		
<b>Sample Size<sup>2</sup></b>	households	persons (thereof children)	
	NET	4,528	11,366 (2,290)
	GROSS	7,430	
<b>Field Period</b>	February to October 1984		
<b>Initial Survey Mode</b>	Paper-and-Pencil Interivewing (PAPI), possibility for self completion		
<b>Number of Interviewers</b>	592		
<b>Initial Response Rate<sup>3</sup></b>	60.9%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	5,491	1,402	3,358 / 11,041
<b>Further Readings</b>			

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

---

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

<sup>2</sup>The 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).

<sup>3</sup>AAPOR Response Rate Definition RR2, see AAPOR (2016).

## 2.2 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

---

<b>Sampling Design</b>	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		
<b>Sample Size</b>	households	persons (thereof children)	
	NET	1,393	4,807 (1,638)
	GROSS	2,045	
<b>Field Period</b>	April to October 1984		
<b>Initial Survey Mode</b>	Paper-and-Pencil Interivewing (PAPI)		
<b>Number of Interviewers</b>	253		
<b>Initial Response Rate</b>	68.1%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	820	574	89 / 4,347
<b>Further Readings</b>	<p><a href="#">Infratest Sozialforschung (2011). SOEP 1984 – Methodenbericht zum Befragungsjahr 1984 (Welle 1) des Sozio-oekonomischen Panels. SOEP Survey Papers 1, DIW/SOEP, Berlin 2011.</a></p>		

## 2.3 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

---

<b>Sampling Design</b>	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 <sup>4</sup>		
<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		
<b>Sample Size</b>	households	persons (thereof children)	
	NET	2,179	6,044 (1,591)
	GROSS	3,404	
<b>Field Period</b>	May to July 1990		
<b>Initial Survey Mode</b>	Paper-and-Pencil Interivewing (PAPI), possibility for self completion		
<b>Number of Interviewers</b>	215		
<b>Initial Response Rate</b>	64.0%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	3,103	1,734	367 / 19,102
<b>Further Readings</b>	<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>		

---

<sup>4</sup>In German: Abteilung Soziologische Forschung des Rundfunks der DDR.

## 2.4 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 which 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 <sup>5</sup>	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 Interivewing (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)	3,905	1,716	1,699 / 9,853

#### 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.

---

<sup>5</sup>213 cases in Sample D do not meet the requirements of the SOEP sampling design. These cases are interviewed, but do not receive valid weights.

## 2.5 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 Interview (CAPI) was implemented. Interviews in Samples A-D at this time were completely conducted using Paper-and-Pencil-Interviews (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

---

<b>Sampling Design</b>	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		
<b>Sample Size</b>	households	persons (thereof children)	
	NET	1,056	2,376 (466)
	GROSS	1,969	
<b>Field Period</b>	April to September 1998		
<b>Initial Survey Mode</b>	Computer assisted personal interview (CAPI) and Paper-and-Pencil-Interviewing (PAPI)		
<b>Number of Interviewers</b>	130		
<b>Initial Response Rate</b>	53.6%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	35,568	18,294	14,827 / 205,099

### Further Readings

Infratest Sozialforschung (2011). *SOEP 1998 – Methodenbericht Erstbefragung der Stichprobe E zum Befragungsjahr 1998 (Welle 15) des Sozio-ökonomischen Panels*. SOEP Survey Papers 33, DIW/SOEP, Berlin 2011.

Projektgruppe Das Sozio-ökonomische Panel (DIW) (1998). *Funktion und Design einer Ergänzungstichprobe für das Sozio-ökonomische Panel (SOEP)*. DIW Discussion Papers 163, Berlin 1998.

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.

## 2.6 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

---

<b>Sampling Design</b>	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		
<b>Sample Size</b>	households	persons (thereof children)	
	NET	6,043	13,871 (2,991)
	GROSS	11,862	
<b>Field Period</b>	March to October 2000		
<b>Initial Survey Mode</b>	Computer assisted personal interview (CAPI) and personal oral interview with possibility for self completion (PAPI)		
<b>Number of Interviewers</b>	671		
<b>Initial Response Rate</b>	50,9%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	6,364	2,224	2,376 / 18,861
<b>Further Readings</b>	<p><a href="#">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.</a></p>		

## 2.7 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

---

<b>Sampling Design</b>	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>	Statification according income and region (east/west)	
<b>Sample Size</b>		households	persons (thereof children)
	NET	1,224	3,364 (693)
	GROSS	2,493	
<b>Field Period</b>	March to July 2002		
<b>Initial Survey Mode</b>	Computer assisted personal interview (CAPI) and personal oral interview with possibility for self completion (PAPI)		
<b>Number of Interviewers</b>	276		
<b>Initial Response Rate</b>	49,0%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	2,084	953	983 / 9,757
<b>Further Readings</b>			

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.



## 2.8 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

---

<b>Sampling Design</b>	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		
<b>Sample Size</b>	households	persons (thereof children)	
	NET	1,506	3,239 (623)
	GROSS	3,750	
<b>Field Period</b>	March to July 2006		
<b>Initial Survey Mode</b>	Computer assisted personal interview (CAPI)		
<b>Number of Interviewers</b>	243		
<b>Initial Response Rate</b>	40,0%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	26,443	13,454	9,024 / 128,852
<b>Further Readings</b>	<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>		

## 2.9 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

---

<b>Sampling Design</b>	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	
<b>Sample Size</b>	households	persons (thereof children)
	NET	1,495      3,052 (620)
	GROSS	4,743
<b>Field Period</b>	September 2009 to January 2010	
<b>Initial Survey Mode</b>	Computer Assisted Personal Interview (CAPI)	
<b>Number of Interviewers</b>	233	
<b>Initial Response Rate</b>	31.5%	
<b>Further Readings</b>		

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. Hajeck, 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.

## 2.10 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

---

<b>Sampling Design</b>	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 <sup>6</sup> such that the share of migrants for each PSU is doubled		
	Selected unit: household		
<b>Sample Size</b>		households	persons (thereof children)
	NET	3,136	6,308 (1,147)
	GROSS	9,492	
<b>Field Period</b>	March to October 2011		
<b>Initial Survey Mode</b>	Computer Assisted Personal Interview (CAPI)		
<b>Number of Interviewers</b>	338		
<b>Initial Response Rate</b>	33.0%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	12,590	6,191	1,797 / 49,580

### 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.

---

<sup>6</sup>Identification of potentially migrant households using onomastic procedure.

## 2.11 Sample K (2012)

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

### Key Facts

---

<b>Sampling Design</b>	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		
<b>Sample Size</b>	households	persons (thereof children)	
	NET	1,526	3,036 (563)
	GROSS	4,397	
<b>Field Period</b>	March to October 2012		
<b>Initial Survey Mode</b>	Computer assisted personal interview (CAPI)		
<b>Number of Interviewers</b>	304		
<b>Initial Response Rate</b>	34.7%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	26,053	10,204	3,723 / 81,563

### 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.

## 2.12 Sample L1 (FiD) (2010)

Sample L1 “Cohort Sample”<sup>7</sup> 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

<b>Sampling Design</b>	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 <sup>8</sup> such that the share of migrants for each PSU is doubled		
	Selected unit: child in the respective cohort		
<b>Sample Size</b>	households	persons (thereof children)	
	NET	2,074	7,670 (3,900)
	GROSS	5,286	
<b>Field Period</b>	June to October 2010		
<b>Initial Survey Mode</b>	Computer assisted personal interview (CAPI)		
<b>Number of Interviewers</b>	204		
<b>Initial Response Rate</b>	39.2%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	935	577	74 / 3,497
<b>Further Readings</b>	<p><a href="#">TNS Infratest Sozialforschung (2010). "Familien in Deutschland" (FiD) 2010 Methodenbericht: Anlage und Ergebnisse der FiD-Stichproben. München 2011.</a></p> <p><a href="#">Schröder, M., R. Siegers, K. Spieß (2013). "Familien in Deutschland" - FiD. Schmollers Jahrbuch: Vol. 133, No. 4, pp. 595-606.</a></p>		

<sup>7</sup>Sample 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.

<sup>8</sup>Identification of potentially migrant addresses using onomastic procedure and information on the citizenship.

## 2.13 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

---

<b>Sampling Design</b>	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		
<b>Sample Size</b>	households	persons (thereof children)	
	NET	2,500 <sup>9</sup>	8,838 (4,611)
	GROSS	3,281	
<b>Field Period</b>	March to June 2010		
<b>Initial Survey Mode</b>	Computer assisted personal interview (CAPI)		
<b>Number of Interviewers</b>	343		
<b>Initial Response Rate</b>	76.2%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	1,596	1,035	213 / 7,701
<b>Further Readings</b>	<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>		

---

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

## 2.14 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

---

<b>Sampling Design</b>	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		
<b>Sample Size</b>	households	persons (thereof children)	
	NET	924 <sup>10</sup>	3,579 (2,092)
	GROSS	1,144	
<b>Field Period</b>	March to June 2011		
<b>Initial Survey Mode</b>	Computer assisted personal interview (CAPI)		
<b>Number of Interviewers</b>	250		
<b>Initial Response Rate</b>	80.8%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	2,359	1,582	468 / 12,146
<b>Further Readings</b>			

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.

---

<sup>10</sup>During 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.

## 2.15 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

<b>Sampling Design</b>	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 <sup>11</sup> 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		
<b>Sample Size</b>	households	persons (thereof children)	
	NET	2,732	7,445 (2,481)
	GROSS	12,196	
<b>Field Period</b>	May to November 2013		
<b>Initial Survey Mode</b>	Computer assisted personal interview (CAPI)		
<b>Number of Interviewers</b>	232		
<b>Initial Response Rate</b>	31,7% <sup>12</sup>		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	1,564	1,526	64 / 9,225
<b>Further Readings</b>	<p>TNS Infratest Sozialforschung (2014). <i>Methodenbericht zum IAB-SOEP-Migrationssample 2013</i>. SOEP Survey Papers 217, DIW/SOEP, Berlin 2014.</p> <p>Kroh, M., S. Kühne, J. Goebel and F. Preu (2015). <i>The 2013 IAB-SOEP Migration Sample (M1): Sampling Design and Weighting Adjustment</i>. SOEP Survey Papers 271, DIW/SOEP, Berlin 2015.</p> <p>Eisnecker, P. S., K. Erhardt, M. Kroh, and P. Trübswetter (2017). <i>The Request for Record Linkage in the IAB-SOEP Migration Sample</i>. SOEP Survey Papers 291, DIW/SOEP, Berlin 2017.</p> <p>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: <i>Public Opinion Quarterly</i> 81.1, 131-143</p>		

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

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



## 2.16 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

---

<b>Sampling Design</b>	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 <sup>13</sup> in each stratum		
<i>second stage</i>	Disproportional address sampling in each PSU according to country of origin		
	Selected unit: person		
<b>Sample Size</b>	households	persons (thereof children)	
	NET	1,096	2,638 (927)
	GROSS	6,813	
<b>Field Period</b>	May to December 2015		
<b>Initial Survey Mode</b>	Computer assisted personal interview (CAPI)		
<b>Number of Interviewers</b>	143		
<b>Initial Response Rate</b>	28,8% <sup>14</sup>		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	899	744	51 / 3,390
<b>Further Readings</b>	<p><a href="#">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.</a></p>		

<sup>13</sup>Identified by the year they entered the IEB and former and current citizenship.

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

## 2.17 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

<b>Sampling Design</b>	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		
<b>Sample Size</b>	households	persons (thereof children)	
	NET	3,336	9,965 (5,438)
	GROSS	6,848	
<b>Field Period</b>	June to December 2016		
<b>Initial Survey Mode</b>	Computer assisted personal interview (CAPI)		
<b>Number of Interviewers</b>	164		
<b>Initial Response Rate</b>	48,7%		
<b>Initial Weighting Factor</b>	Average	SD	min / max
	141	197	5 / 3,503
<b>Further Readings</b>			

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., 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.

### 3 Developments in Sample Size

With respect to developments in sample size, the following figures focus on (3.1) comparing the number of successful interviews by cross-section, (3.2) providing a longitudinal study of panel attrition among the original sample members, (3.3) showing the entrance of new sample members by birth / moving into SOEP households and their participation behavior, (3.4) reporting share of original households in relation to new households from splits and (3.5) 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.

#### 3.1 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.

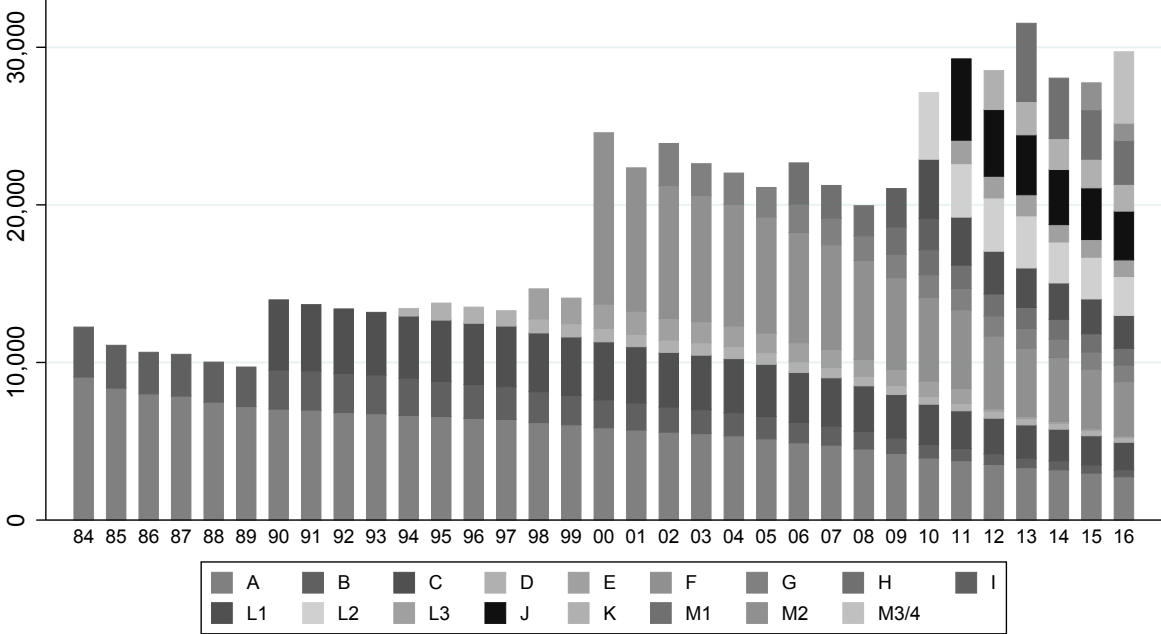


Figure 1: The Number of Successful Interviews with Persons by Subsamples A through M3/4, Waves 1 to 33.

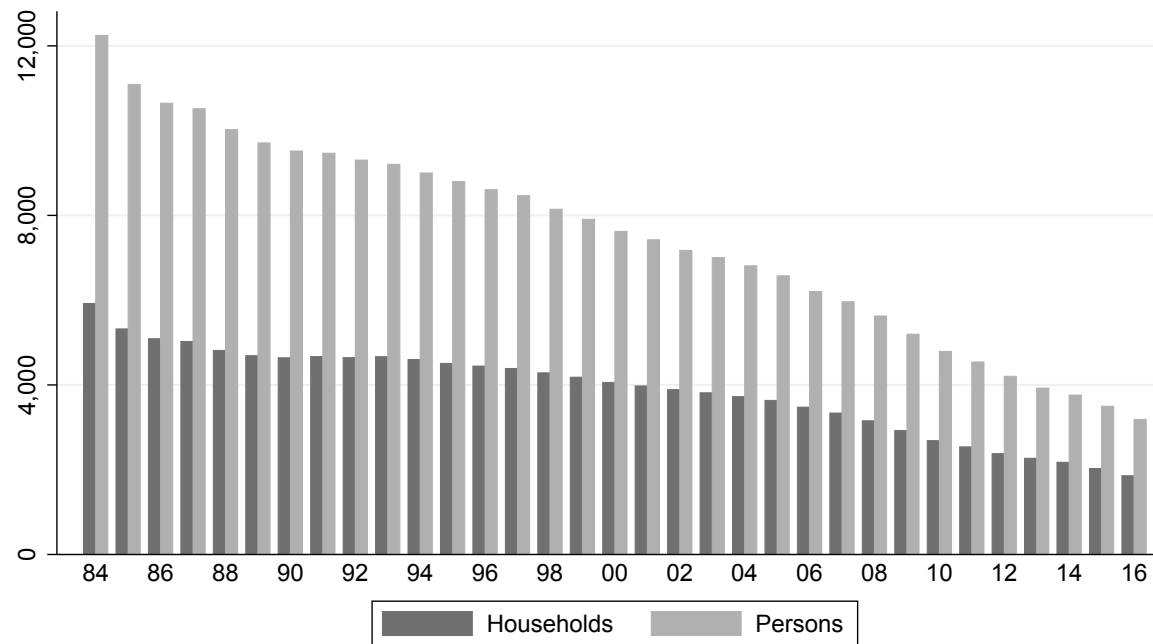


Figure 2: Comparison of Successful Interviews with Persons and Households (Subsamples A and B), Waves 1 to 33

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

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Persons	7,623	7,424	7,175	7,004	6,811	6,575	6,203	5,961	5,626	5,197	4,793	4,541	4,204	3,926	3,761	3,497	3,187
Households	4,060	3,977	3,889	3,814	3,724	3,635	3,476	3,337	3,154	2,923	2,686	2,539	2,379	2,270	2,176	2,028	1,857

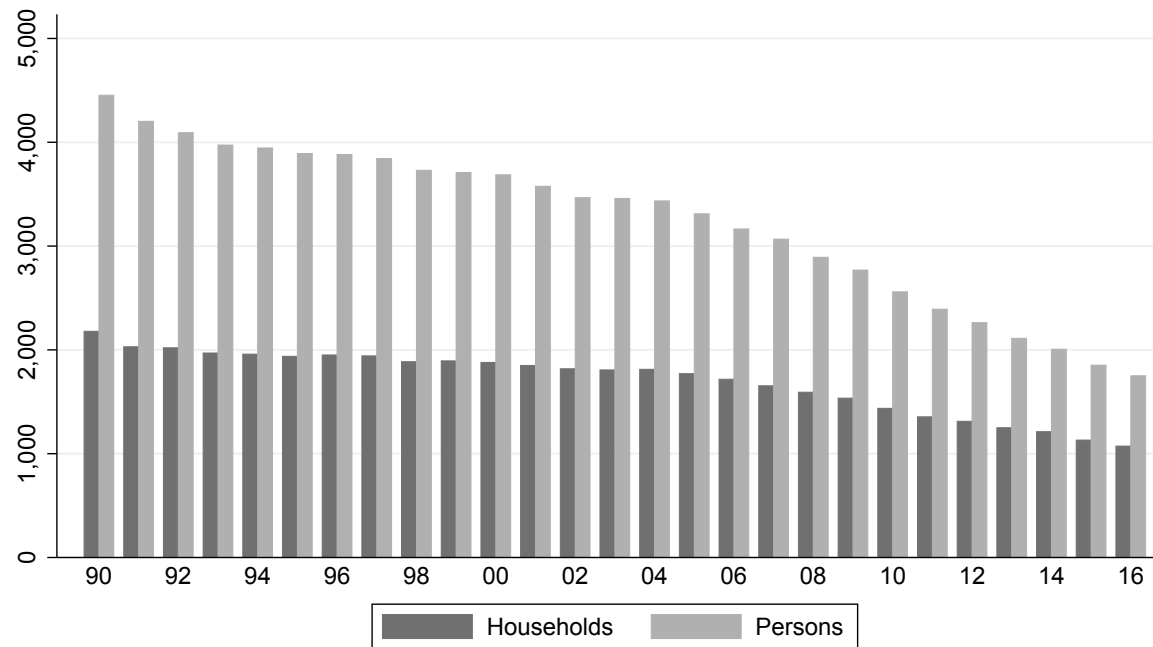


Figure 3: Comparison of Successful Interviews with Persons and Households (Subsamples C), Waves 1 to 27

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
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
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

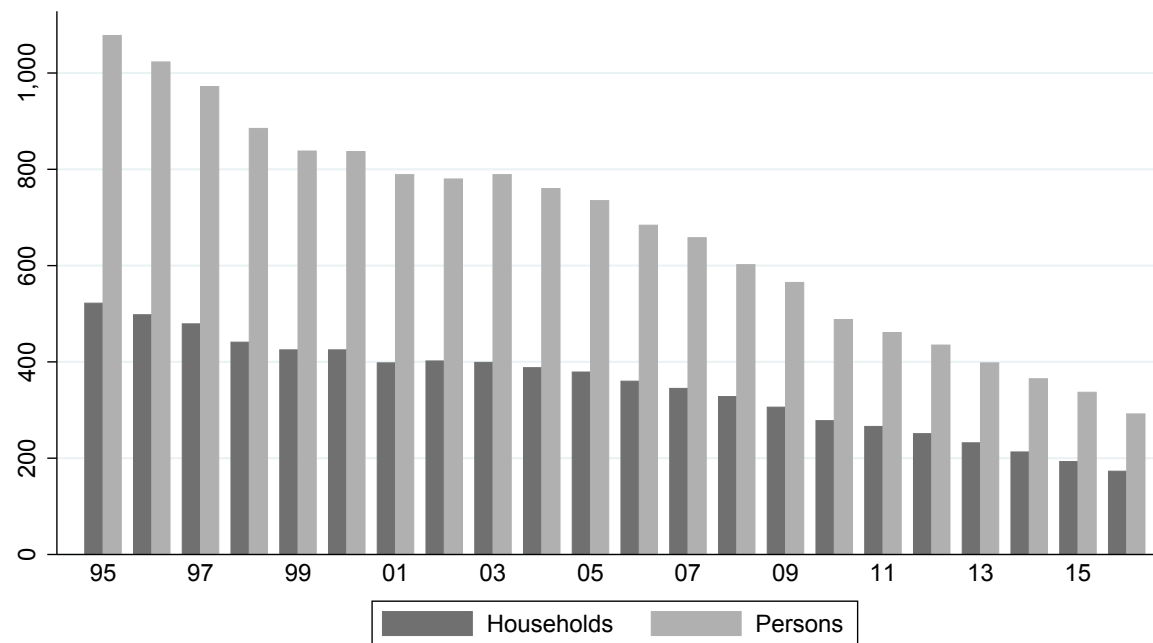


Figure 4: Comparison of Successful Interviews with Persons and Households (Subsamples D), Waves 1 to 22

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

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Persons	684	658	602	565	488	461	435	398	365	337	292
Households	360	345	328	306	278	266	251	232	213	193	173

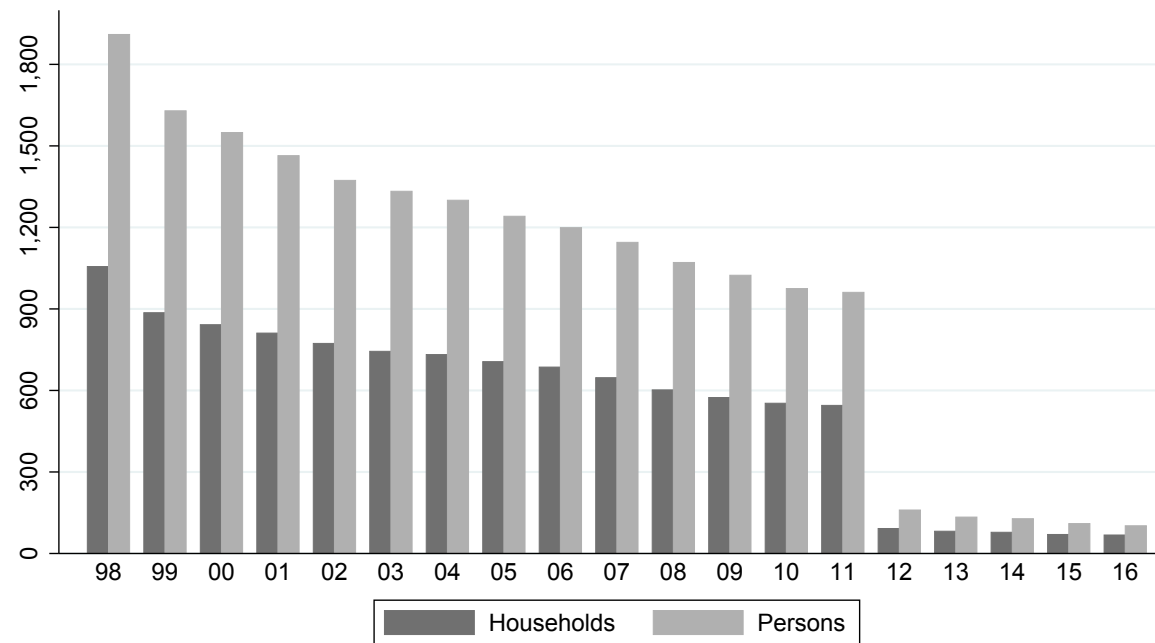


Figure 5: Comparison of Successful Interviews with Persons and Households (Subsample E), Waves 1 to 19<sup>15</sup>

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Persons	1,910	1,629	1,549	1,464	1,373	1,333	1,300	1,241	1,199	1,145	1,071	1,024	975	961	160	134	128	110	102
Households	1,056	886	842	811	773	744	732	706	686	647	602	574	553	545	92	82	78	70	68

<sup>15</sup>In 2012, subsample E has been split into two parts, one being surveyed continuously by SOEP-Core and the larger part being surveyed by SOEP-IS since 2012 onwards.

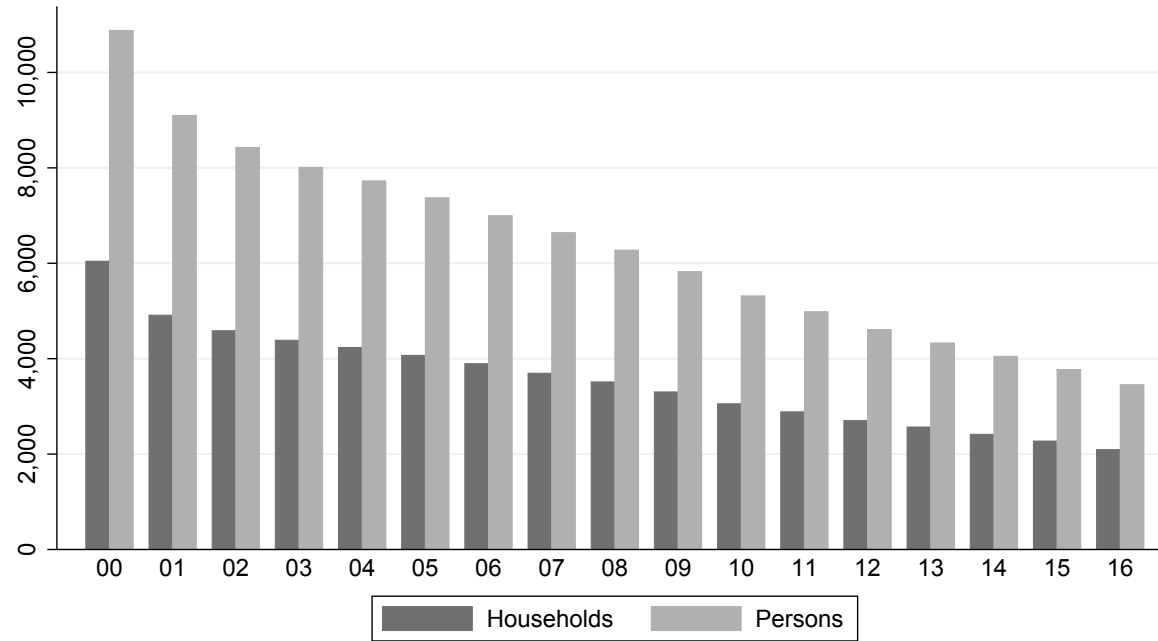


Figure 6: Comparison of Successful Interviews with Persons and Households (Subsample F), Waves 1 to 17

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
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
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



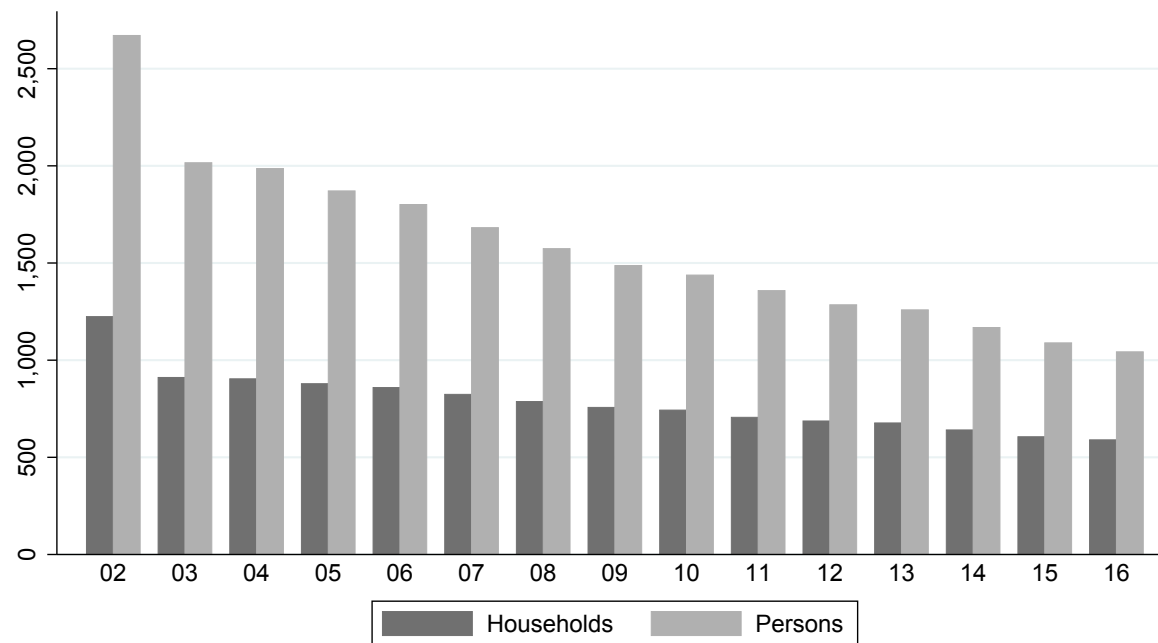


Figure 7: Comparison of Successful Interviews with Persons and Households (Subsample G), Waves 1 to 15<sup>16</sup>

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

<sup>16</sup>In the second wave the target population was changed: a higher income threshold resulted in a smaller number of observations in 2003.

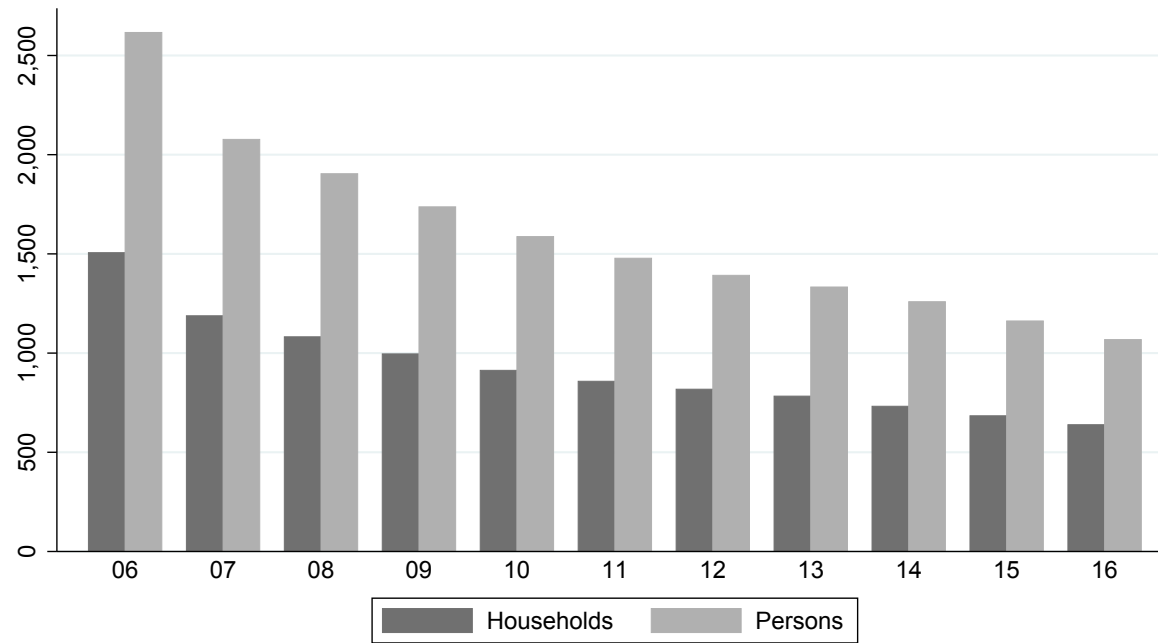


Figure 8: Comparison of Successful Interviews with Persons and Households (Subsample H), Waves 1 to 11

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

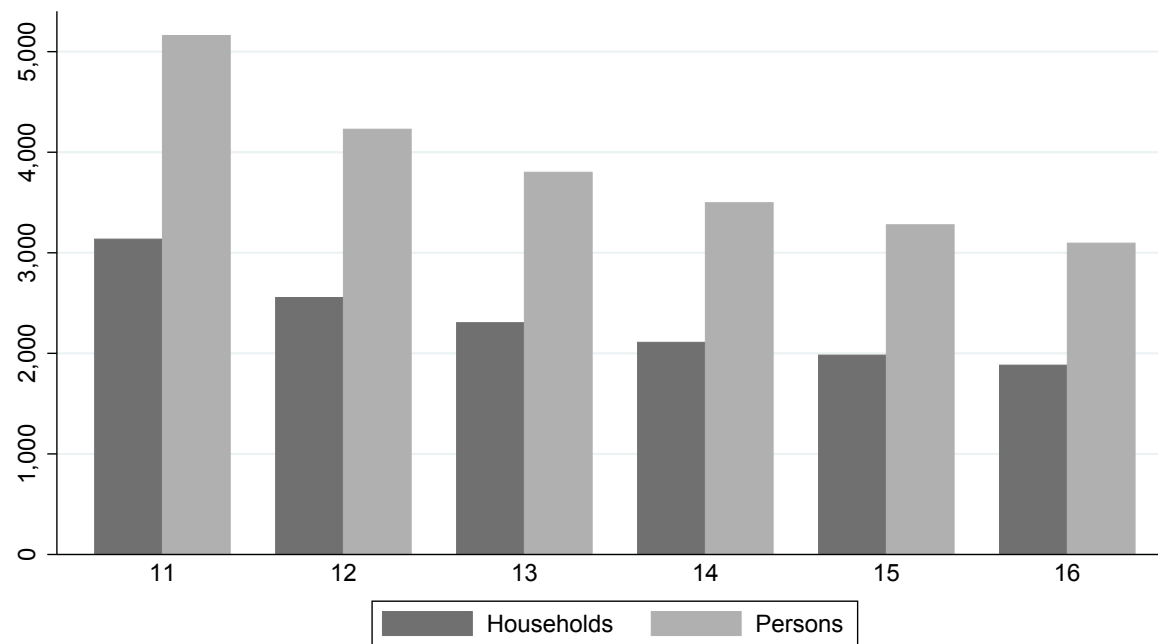


Figure 9: Comparison of Successful Interviews with Persons and Households (Subsample J), Waves 1 to 6

<b>Year</b>	<b>2011</b>	<b>2012</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>	<b>2016</b>
Persons	5,161	4,229	3,801	3,498	3,279	3,096
Households	3,136	2,555	2,305	2,110	1,983	1,883

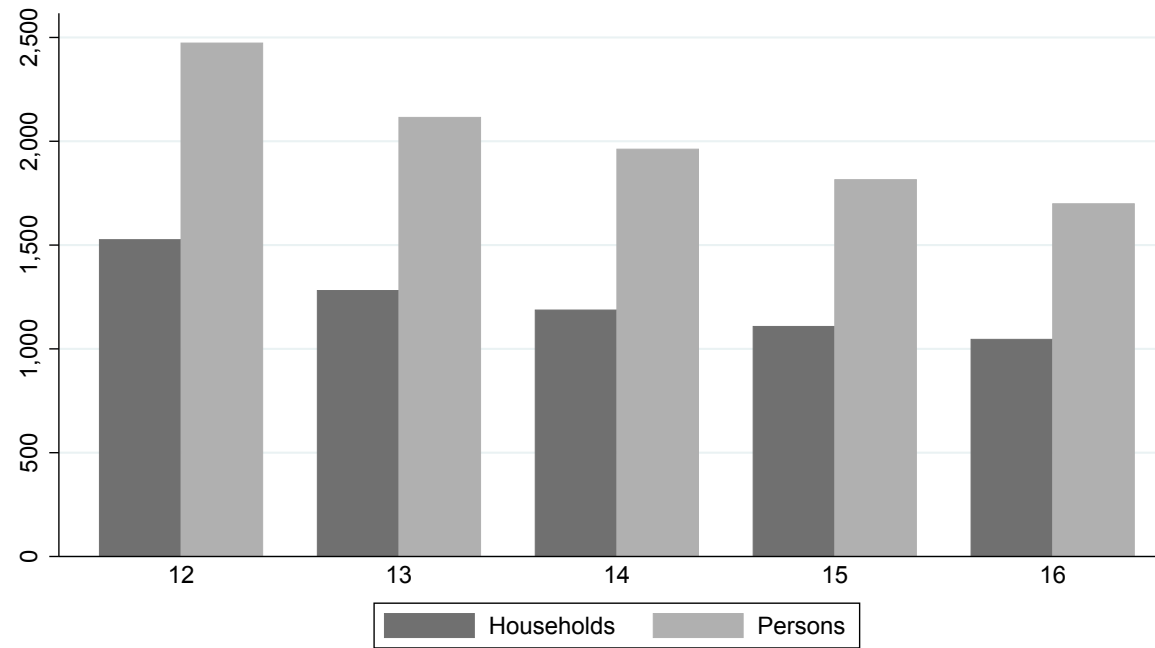


Figure 10: Comparison of Successful Interviews with Persons and Households (Subsample K), Waves 1 to 5

<b>Year</b>	<b>2012</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>	<b>2016</b>
Persons	2,473	2,115	1,962	1,815	1,699
Households	1,256	1,281	1,187	1,108	1,046

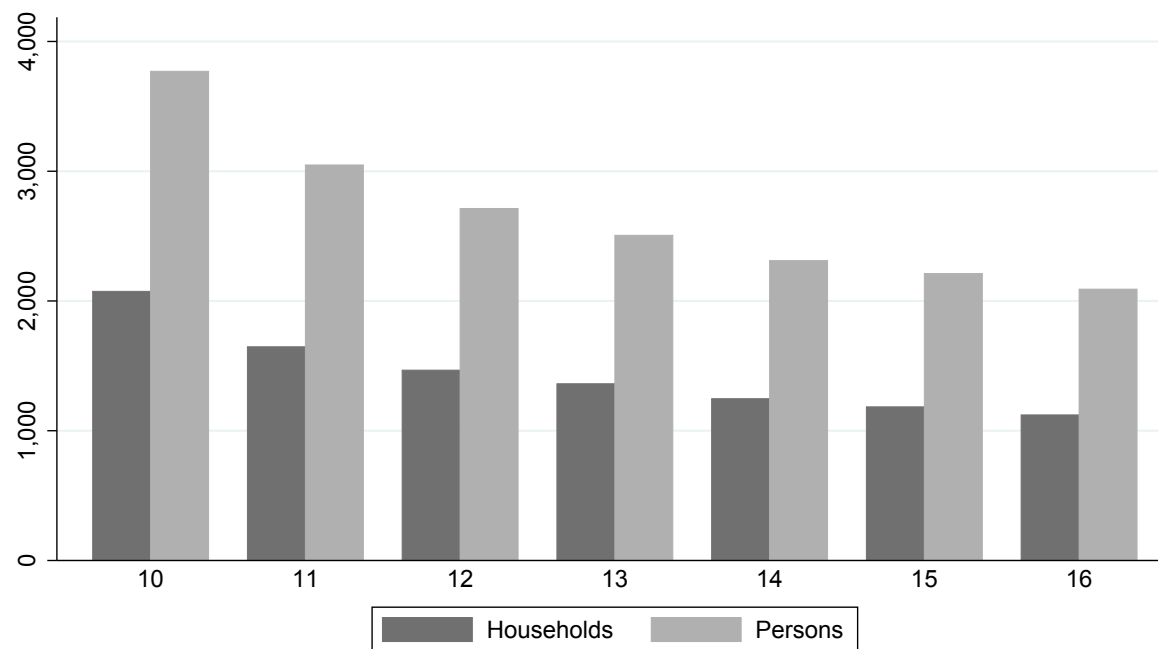


Figure 11: Comparison of Successful Interviews with Persons and Households (Subsample L1), Waves 1 to 7

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

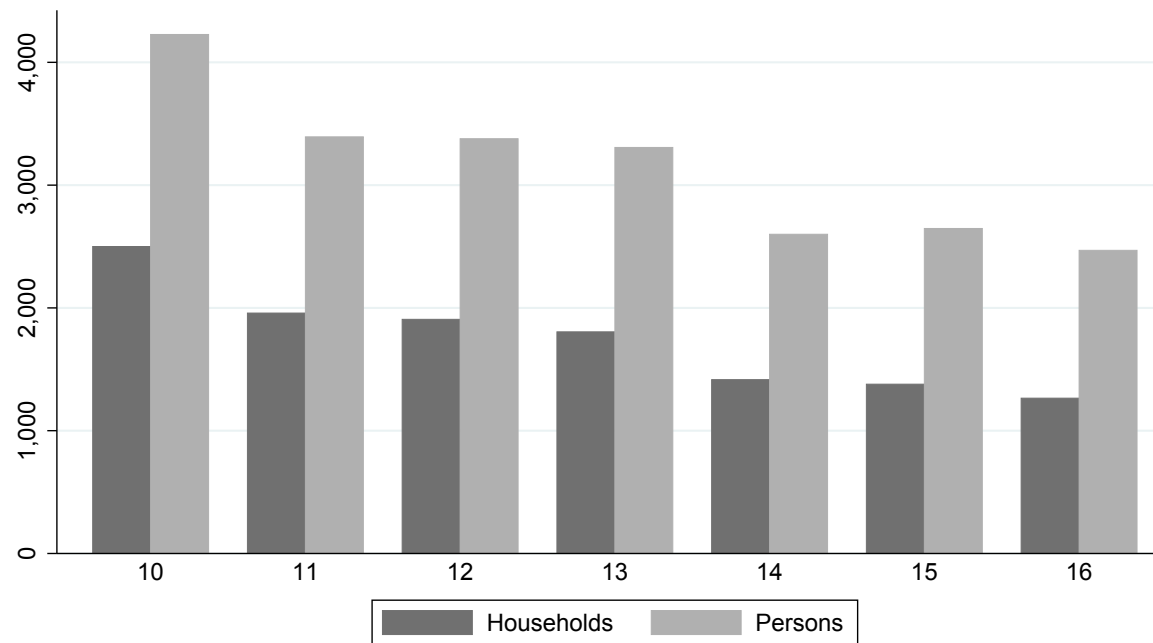


Figure 12: Comparison of Successful Interviews with Persons and Households (Subsample L2), Waves 1 to 7 <sup>17,18</sup>

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

<sup>17</sup>237 households were identified not to be part of the target population and were not followed in the second wave.

<sup>18</sup>In 2014 the default interview mode changed to Computer-Assisted Web Interviewing (CAWI).

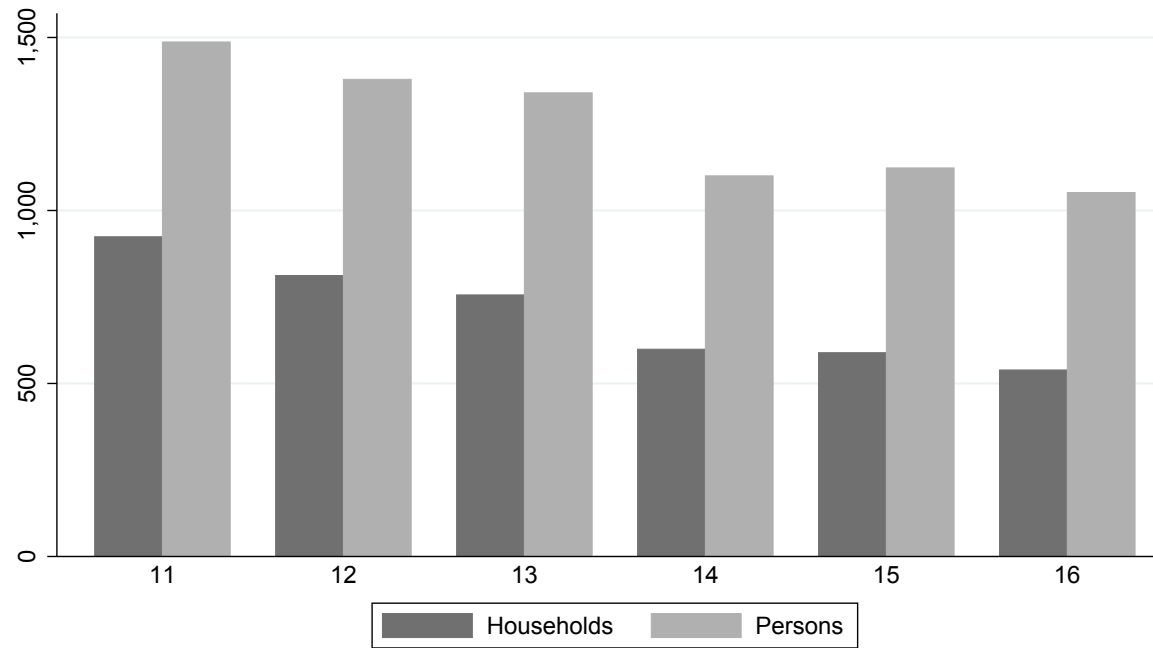


Figure 13: Comparison of Successful Interviews with Persons and Households (Subsample L3), Waves 1 to 6<sup>19</sup>

<b>Year</b>	<b>2011</b>	<b>2012</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>	<b>2016</b>
Persons	1,487	1,379	1,340	1,100	1,123	1,052
Households	924	812	756	599	589	539

<sup>19</sup>In 2014 the default interview mode changed to Computer-Assisted Web Interviewing (CAWI).

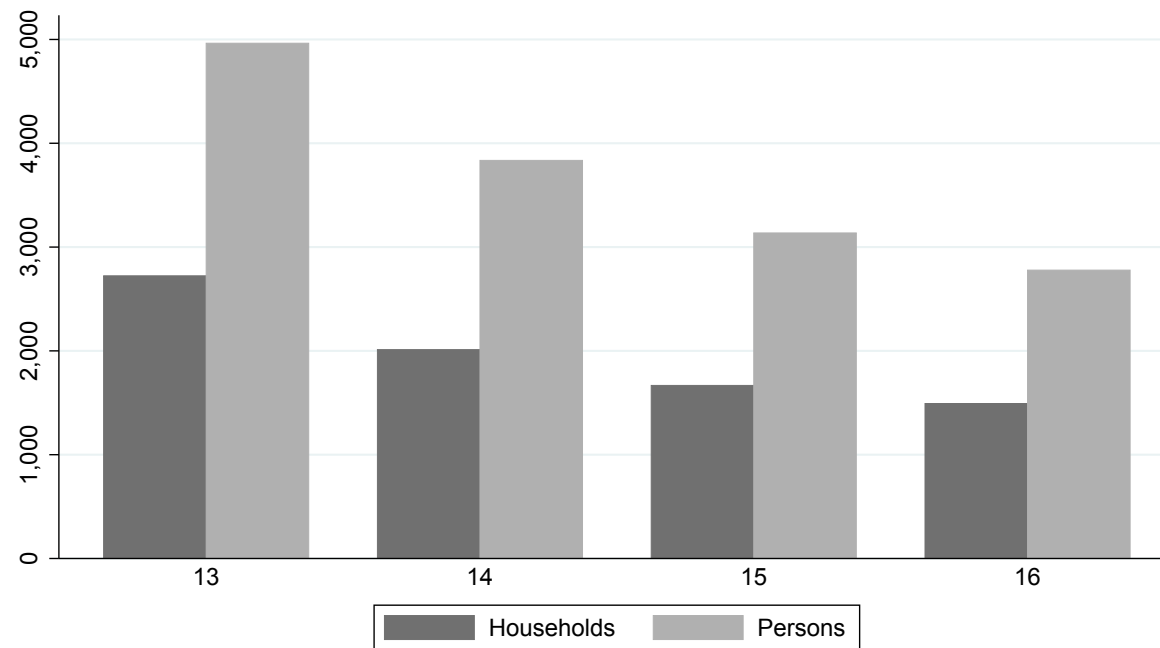


Figure 14: Comparison of Successful Interviews with Persons and Households (Subsample M1), Waves 1 to 4

<b>Year</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>	<b>2016</b>
Persons	4,964	3,835	3,136	2,778
Households	2,723	2,012	1,667	1,493



### 3.2 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”).

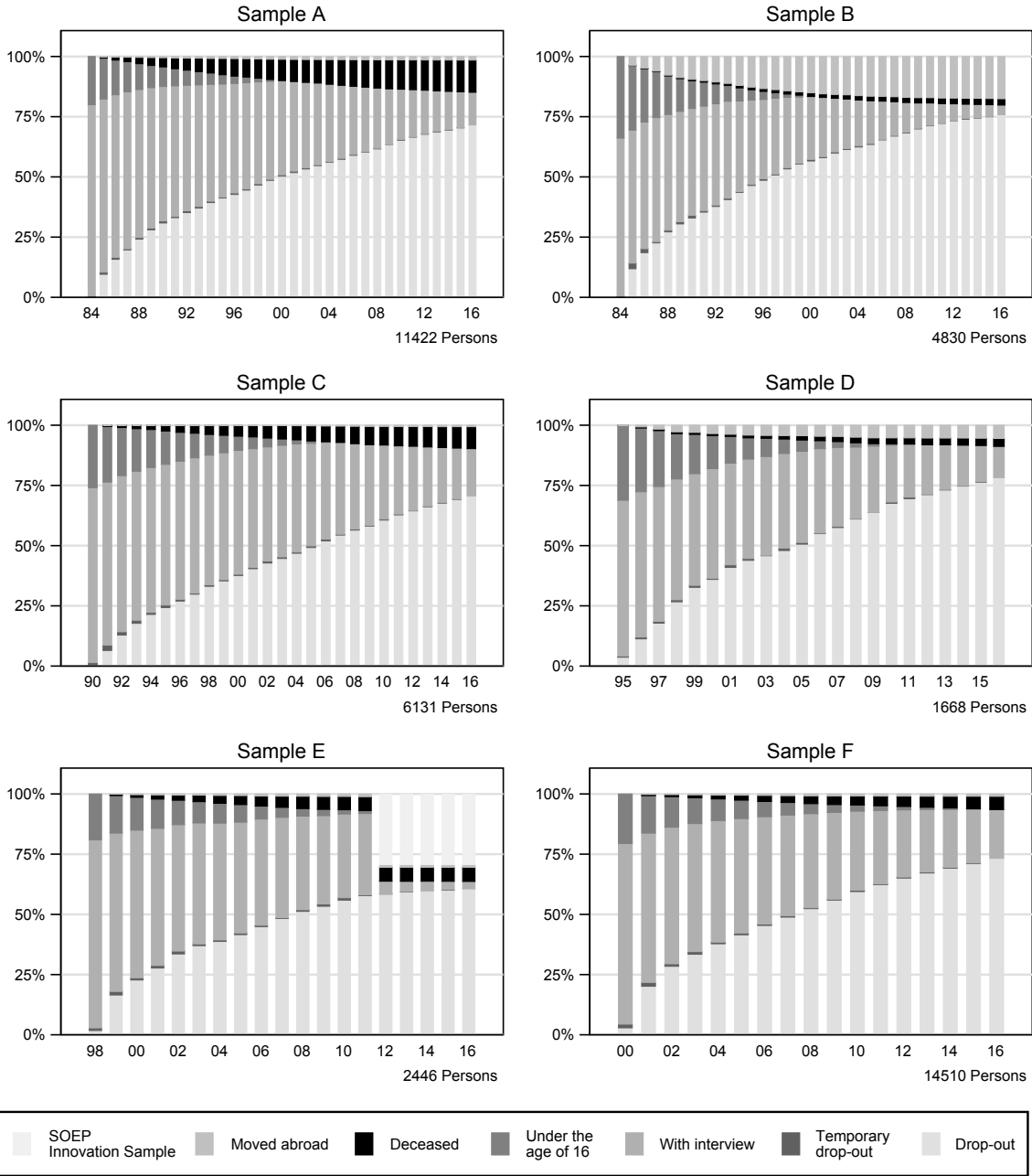


Figure 15: First-Wave Persons and their Participation Behavior. Development up to 2016

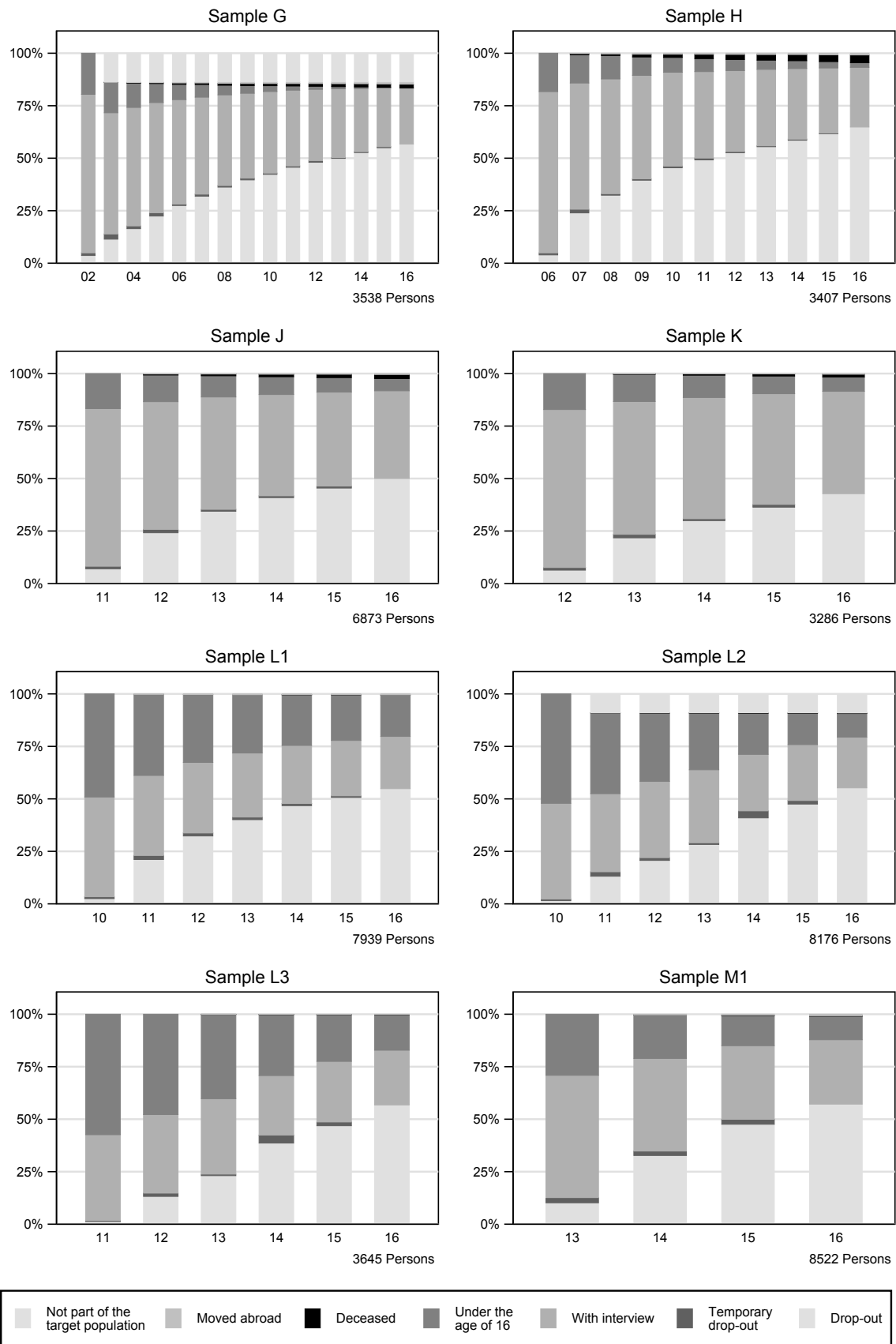


Figure 15: First-Wave Persons and their Participation Behavior. Development up to 2016

### 3.3 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.

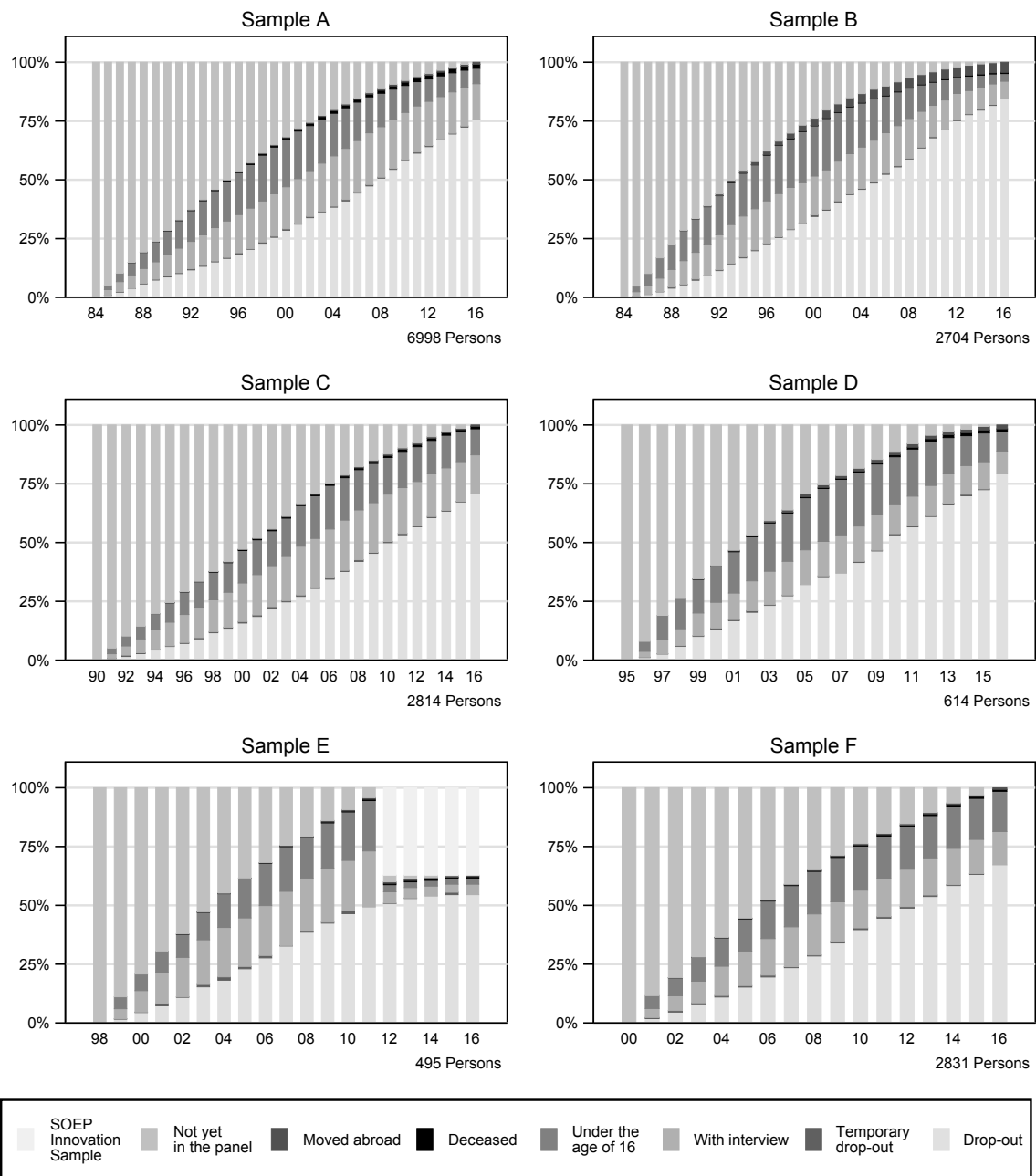


Figure 16: Entrants and their Participation Behavior. Development up to 2016

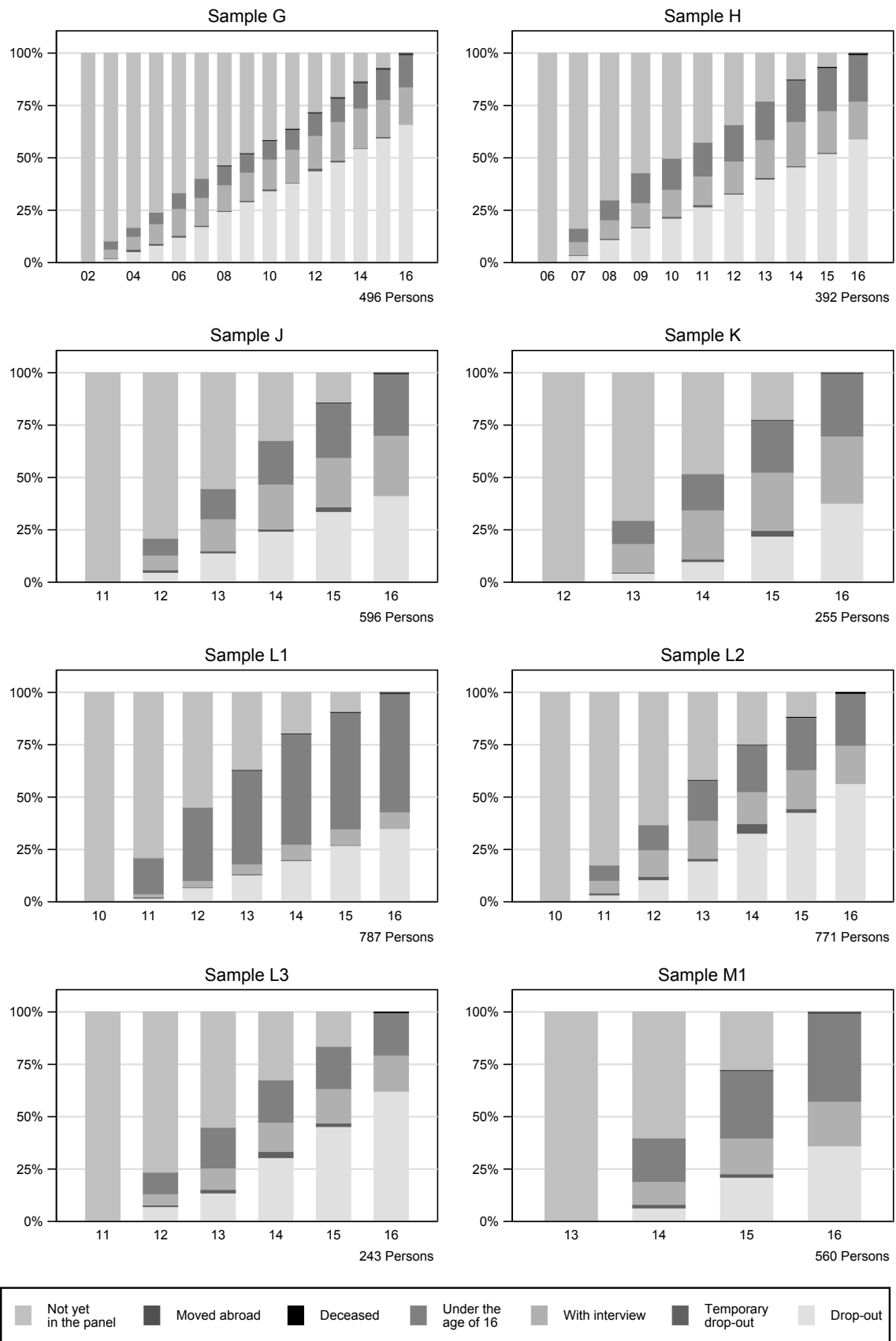


Figure 16: Entrants and their Participation Behavior. Development up to 2016

### 3.4 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 not moving keeps the initial household number. These households are referred to as an “original household”<sup>20</sup>. The following figures display the development of the share of original households for each sample.

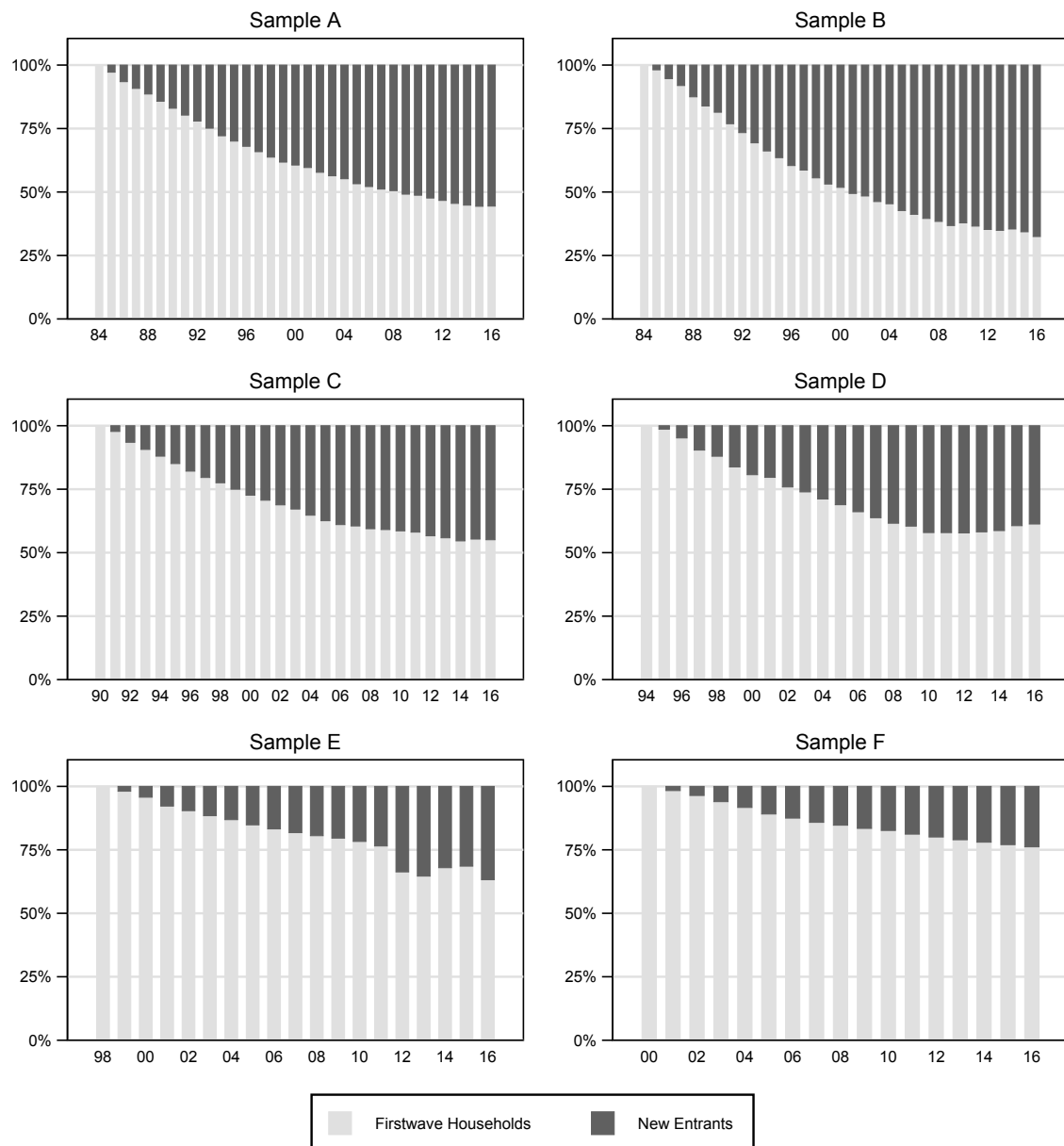


Figure 17: Proportion of First-Wave and New Households. Development up to 2016

<sup>20</sup>For detailed study of the relevance of non-original sample members in the SOEP, see Schonlau et al. (2011) and Spiess et al. (2008).

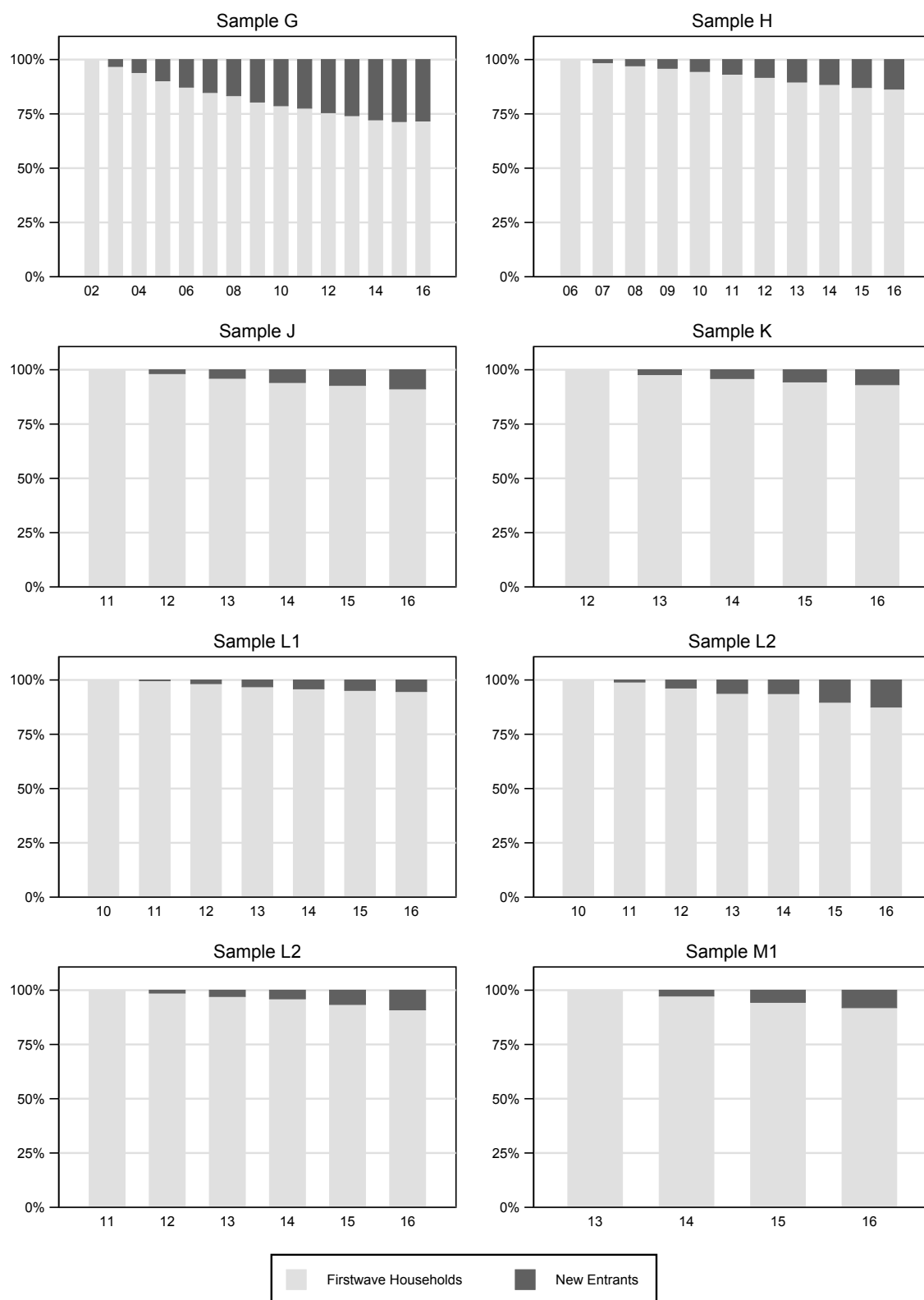


Figure 17: Proportion of First-Wave and New Households. Development up to 2016

### 3.5 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 18, 19, 20 and 21) and some basic socio-demographic characteristics measured in the year of sampling, such as age, occupation, income, and education (Figures 22, 23, 24 and 25). 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 18), 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.

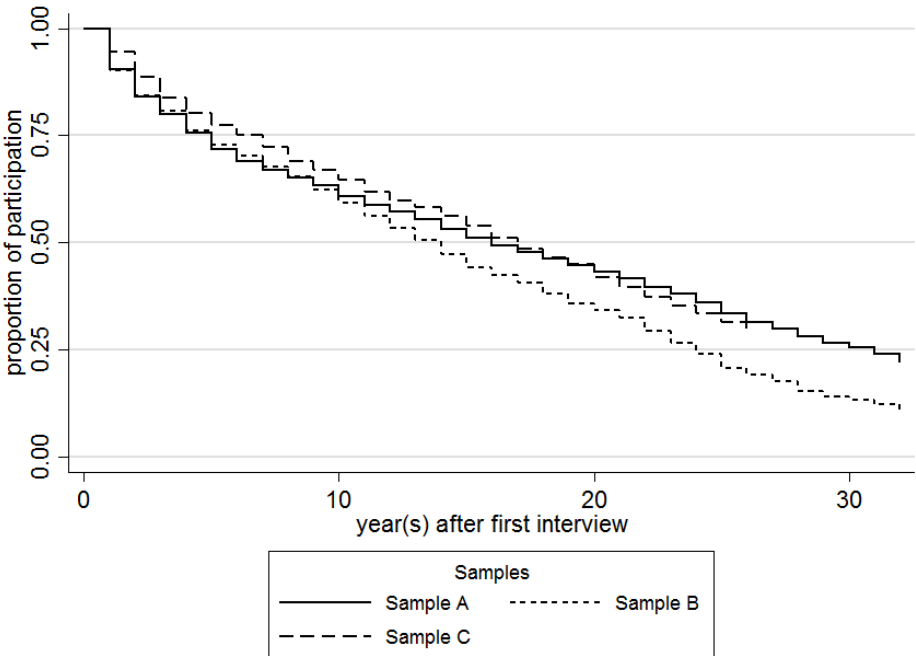


Figure 18: Successful Re-Interviewing of First-Wave Respondents by Subsamples A, B, C. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

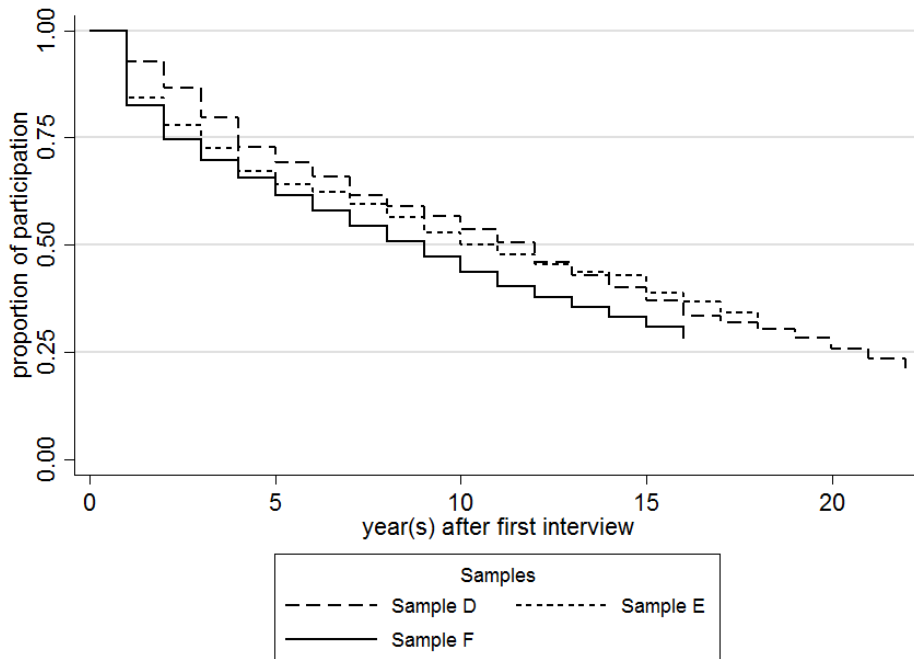


Figure 19: Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

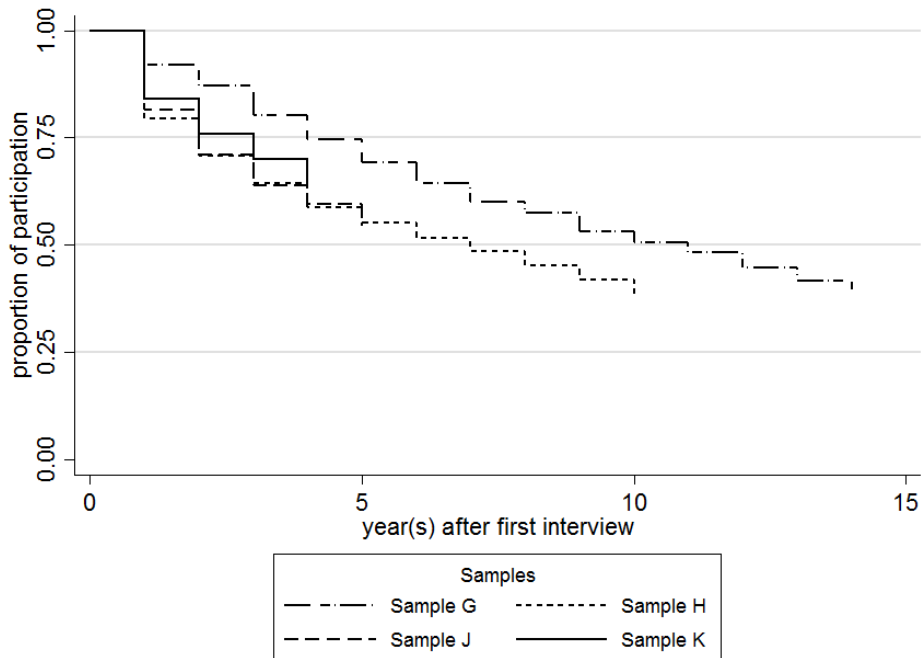


Figure 20: Successful Re-Interviewing of First-Wave Respondents by Subsamples G, H, J and K. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad



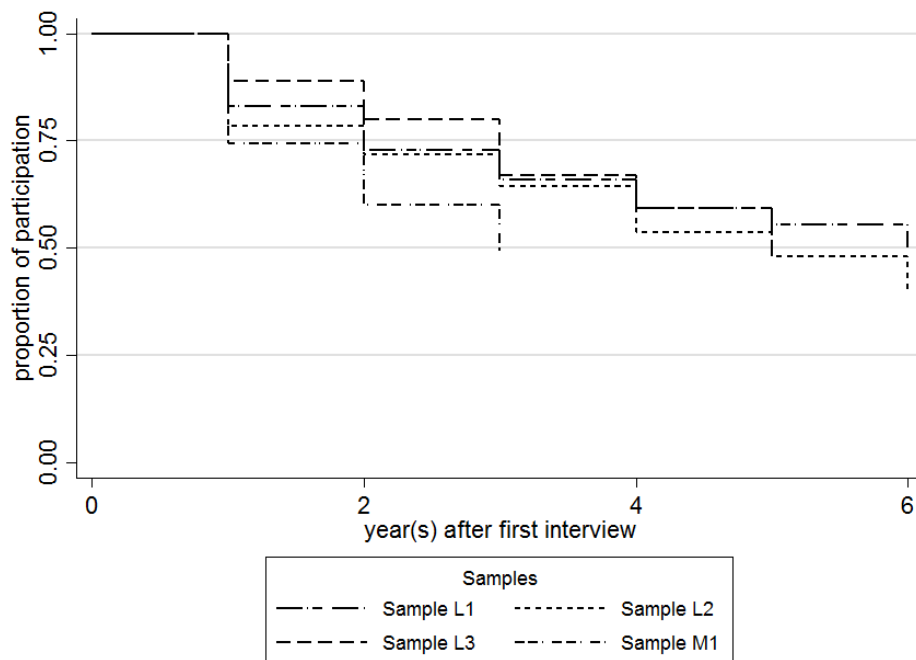


Figure 21: Successful Re-Interviewing of First-Wave Respondents by Subsamples L1, L2, L3 and M1. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

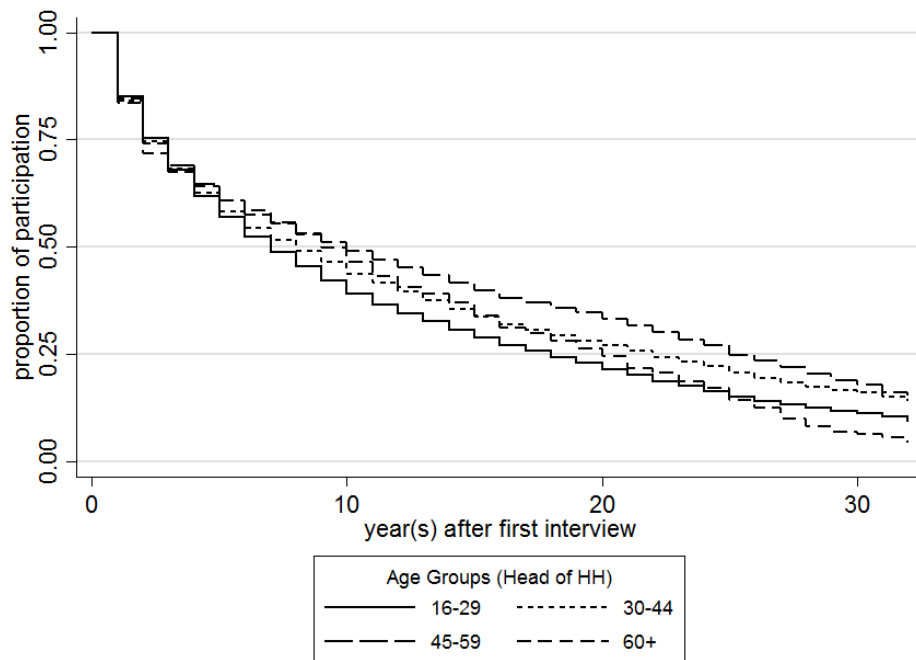


Figure 22: Successful Re-Interviewing of All First-Wave Respondents by Age Categories. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

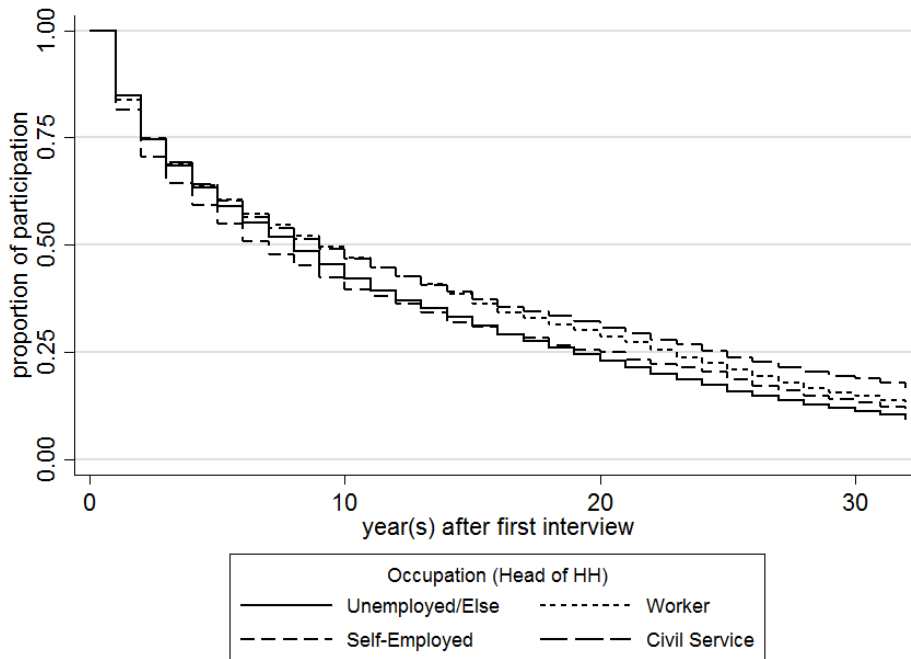


Figure 23: Successful Re-Interviewing of All First-Wave Respondents by Occupation. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

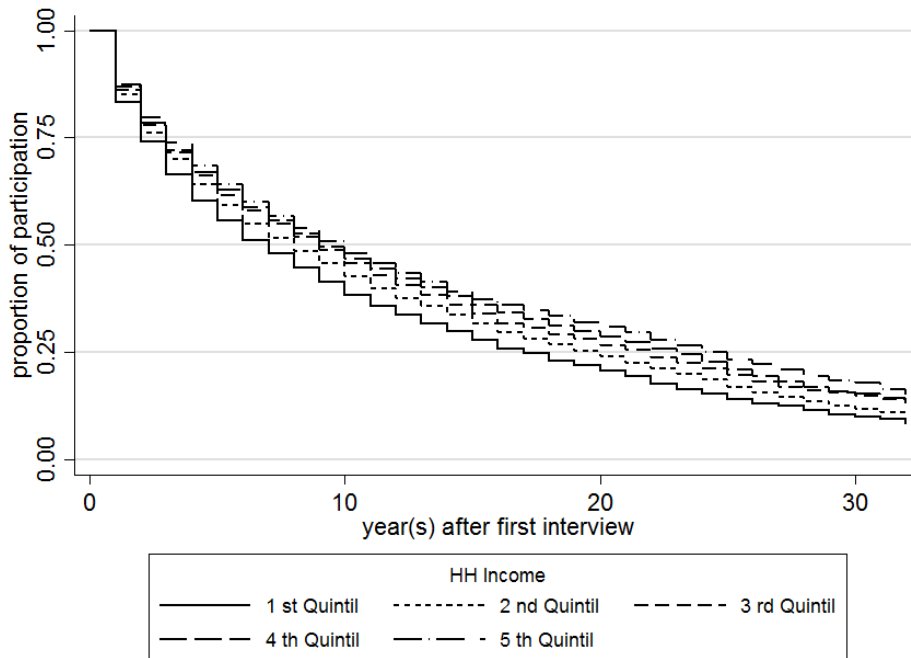


Figure 24: Successful Re-Interviewing of All First-Wave Respondents by Income Quintiles. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

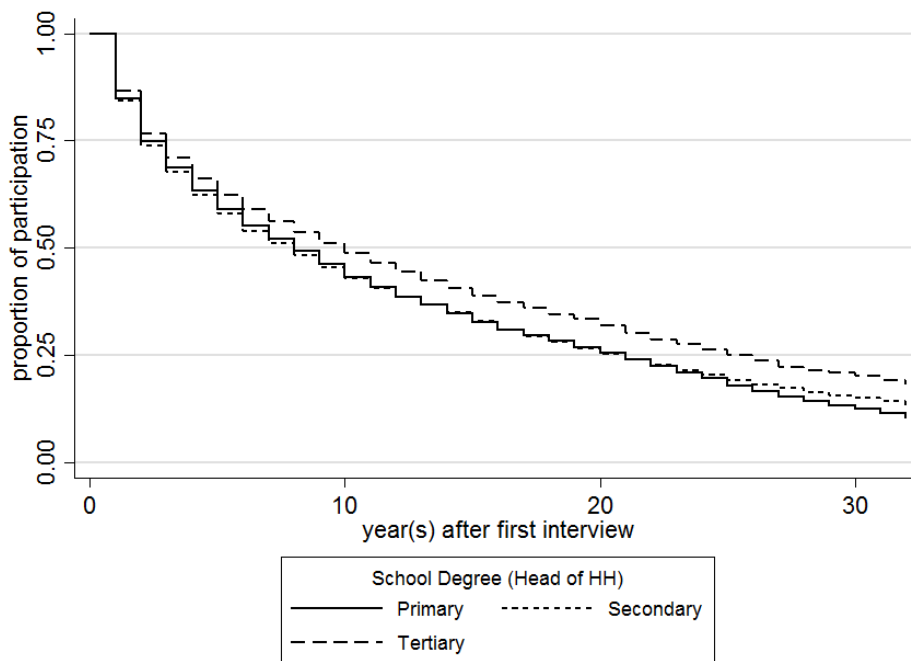


Figure 25: Successful Re-Interviewing of All First-Wave Respondents by Education. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

## 4 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.

### 4.1 The Frequency of Successful Follow-Ups

Table 4.1 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 2016. 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 4.1: The Frequency of Households to be Re-Contacted and the Percentage of Successful Follow-Ups, Subsamples A to M3/4 by Year.

Year	Sample A		Sample B		Sample C		Sample D		Sample E		Sample F		Sample G		Sample H		Sample I		Sample J		Sample K		Sample L1		Sample L2		Sample L3		Sample M1		Sample M2		Sample M3/4	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	
1984	4,528		1,393																															
1985	4,681	98.3	1,370	96.4																														
1986	4,486	98.9	1,325	97.0																														
1987	4,232	99.0	1,220	98.6																														
1988	4,140	99.1	1,191	99.0																														
1989	3,984	99.0	1,157	99.0																														
1990	3,902	99.1	1,124	98.8	2,179																													
1991	3,860	99.5	1,151	99.2	2,246	98.4																												
1992	3,845	99.7	1,153	99.2	2,302	99.4																												
1993	3,867	99.2	1,172	98.5	2,227	99.0																												
1994	3,849	99.2	1,150	98.9	2,134	99.4	236																											
1995	3,784	99.5	1,108	98.9	2,110	99.5	540																											
1996	3,747	99.6	1,069	99.2	2,103	99.4	544	99.6																										
1997	3,688	99.6	1,038	99.0	2,087	99.4	541	99.2																										
1998	3,667	99.4	1,019	99.3	2,079	99.3	528	98.9	1,056																									
1999	3,631	99.6	975	99.3	2,037	99.6	498	99.3	1,089	99.4																								
2000	3,549	99.6	934	99.4	2,025	99.6	467	99.8	967	99.1	6,043																							
2001	3,463	99.5	904	99.4	2,034	99.7	454	99.0	921	99.0	6,162	98.8																						
2002	3,406	99.7	877	99.0	2,005	99.5	450	99.8	873	99.4	5,447	99.4	1,224																					
2003	3,330	99.6	840	99.6	1,982	99.6	434	99.5	834	99.2	4,965	99.7	1,056	99.0																				
2004	3,260	99.8	803	99.6	1,962	99.6	436	99.7	797	99.7	4,736	99.6	1,010	99.7																				
2005	3,220	99.8	779	99.3	1,959	99.7	429	99.2	783	99.9	4,577	99.7	1,001	99.7																				
2006	3,138	99.7	770	99.5	1,941	99.4	425	98.6	775	99.0	4,401	99.2	995	99.4	1,506																			
2007	3,000	99.7	725	99.4	1,834	99.9	387	99.4	727	99.7	4,157	99.5	933	99.2	1,530	99.4																		
2008	2,856	99.8	676	99.1	1,767	99.5	372	99.4	680	99.7	3,962	99.3	904	99.6	1,326	99.5																		
2009	2,730	99.7	620	99.2	1,695	99.9	351	99.7	636	100.0	3,760	99.6	870	99.5	1,145	99.7	1,495																	
2010	2,570	99.8	548	99.3	1,627	100.0	334	99.6	605	99.8	3,538	99.5	826	99.9	1,059	99.5	1,737	97.6					2,074	2,500										
2011	2,421	99.8	495	99.0	1,541	99.8	302	99.3	589	100.0	3,318	99.6	797	99.6	992	99.5			3,136				2,082	98.4	2,271	97.9	924							
2012	2,289	99.8	440	99.7	1,466	99.9	286	100.0	116	98.9	3,076	99.9	774	99.7	928	99.9			3,201	99.1	1,526		1,865	99.5	2,254	98.3	943	98.5						
2013	2,180	99.6	392	99.1	1,417	99.7	269	99.1	98	100.0	2,880	99.7	733	99.6	877	99.5			2,869	99.4	1,564	98.8	1,752	99.1	2,177	98.6	920	99.0	2,723					
2014	2,077	99.3	358	99.3	1,351	99.6	249	100.0	90	100.0	2,741	99.6	725	99.2	828	99.3			2,519	99.0	1,447	99.2	1,510	99.3	2,027	97.5	836	98.0	2,819	98.4				
2015	1,998	99.3	331	99.3	1,300	99.5	229	100.0	83	100.0	2,597	99.1	699	99.2	790	99.7			2,309	99.3	1,360	99.2	1,404	99.3	1,879	97.9	789	97.2	2,453	97.1	1,096			
2016	1,861	99.5	296	99.6	1,217	99.6	208	99.4	83	95.8	2,412	99.3	669	98.5	720	99.5			2,119	99.4	1,209	99.2	1,287	99.4	1,736	98.1	732	97.6	2,111	96.8	1,086	95.4	3,320	

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

## **4.2 Predicting the Probability of Successful vs. Unsuccessful Follow-Ups in the Year 2016**

Based on household and interview level characteristics measured in 2015, we aim at predicting the probability of re-contacting a household relative to unsuccessful follow-up in 2016. 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 4.2 describes the regressors and Table 4.3 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 2015 are not reported in the present data documentation due to space restrictions. These can be obtained from previous attrition documentations.

Table 4.2: Definition of the Regressors of the Logit Model of Unsuccessful Follow-Ups

Variable	Label	Value
<b>Interview Characteristics</b>		
New Household	Household new in SOEP	0/1
New Address	Household moved	0/1
Phone Unknown	Telephone number undisclosed	0/1
Youth Questionnaire	Adolescents in household filled out the youth questionnaire	0/1
Change of Interviewer	Change of interviewer between the two last waves	0/1
<b>Demographic Characteristics</b>		
Single Household	One-person household	0/1
Widowed	Head of household is widowed	0/1
Partner in Household	Head of household lives with partner	0/1
Partner Moved In	Partner of household member moved in two years ago	0/1
Separation	Household member(s) separated from partner	0/1
Naturalized Citizen	Head of household did not acquire german citizenship at birth	0/1
Non-German Nationality	Household member(s) has nationality other than german	0/1
Non-Native Parent	Head of household has parent(s) who is non-native german speaker	0/1
<b>Work, Education and Financial Situation</b>		
In Education	Head of household is currently in education or training	0/1
Blue-Collar Worker	Head of household is a blue-collar worker	0/1
Low Disposable Income	Low disposable income within the 1st quartile	0/1
No Monthly Savings	Household can not put money aside every month	0/1
No Savings for Emergencies	Household has no money put aside for emergencies	0/1
<b>Personality Traits and Well-Being</b>		
Strong Polit. Interest	Head of household is very interested in politics	0/1
No Antiforeignism Concerns	Household member(s) not worried about hostility to foreigners	0/1
Skills Importance	Household member(s) does not agree that skills are more imp. than efforts	0/1
Success through Work	Household member(s) does not agree that one has to work hard to succeed	0/1
Revenge for Injustice	Household member(s) would take a revenge if suffered a great wrong	0/1
Quick recovery	Household member(s) gets over quickly after hurt feelings	0/1
Other Illness	Household member(s) diagnosed with other illness(es)*	0/1
Asthma	Household member(s) diagnosed with asthma	0/1
<b>Building, Area and Region</b>		
Subtenant	Household members are subtenants of the dwelling	0/1
High-Rise Building	Household lives in a high-rise building (9 or more stories)	0/1
SPD Voters Area	Household located in area with high share of SPD voters	0/1

\*Illness(es) other than sleep disorder, diabetes, asthma, cardiac disease, cancer, migraine, high blood pressure, depression, dementia, joint disease

Table 4.3: Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2016

	Sample A	Sample C	Sample F	Sample G	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1	Sample M2
Intercept	3,49***	1,98***	2,07***	2,04***	2,17***	2,77***	3,05***	2,67***	2,43***	1,74***	1,82***
<b>Interview Characteristics</b>											
New Household	-1,70***		-1,19***		-1,25***			-0,60**	-0,97**	-0,91***	-1,40***
New Address					-0,86***	-0,80*	-1,26**	-1,17***	-1,34***	-1,15***	
Phone Unknown						-1,32***		-0,64**	-0,65*	-0,62***	-0,58***
Youth Questionnaire				-1,04**							
Change of Interviewer				-0,83**							
<b>Demographic Characteristics</b>											
Single Household							-1,78***	-0,68**		-0,48***	
Widowed								-0,92*			
Partner in Household											0,38**
Partner Moved In	-1,57**										
Separation		-1,09**									
Naturalized Citizen										0,59**	-0,88**
Non-German Nationality							-0,84*				
Non-Native Parent								-0,57*			
<b>Work, Education and Financial Situation</b>											
In Education	-1,24**	-0,64*		-0,91***							
Blue-Collar Worker	-1,50*										
No Monthly Savings										0,41**	
No Savings for Emergencies											-0,39*
<b>Personality Traits and Well-Being</b>											
Strong Polit. Interest										0,56**	
No Antiforeignism Concerns					-0,39*						
Skills Importance							-0,74*				
Success through Work											-0,85**
Revenge for Injustice			-0,46**								
Quick recovery			-0,34*								
Other Illness									-0,60*		

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

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



Table 4.3 – Continued from previous page

	Sample A	Sample C	Sample F	Sample G	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1	Sample M2
Asthma						-0,77*					
<b>Building, Area and Region</b>											
Subtenant											-0,58**
High-Rise Building						-0,80**					
SPD Voters Area											-0,34*
<i>No. of Observations</i>	1861	1217	2412	669	2119	1209	1287	1736	732	2116	1096
<i>Log Likelihood</i>	-19,20	-18,37	-55,20	-26,49	-42,02	-24,47	-20,62	-68,77	-38,39	-133,40	-96,76

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

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

## **5 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.

### **5.1 The Frequency of Participation**

Table 5.1 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 5.1: The Frequency of Re-Contacted Households and the Percentage of Participation, Subsamples A to M3/4 by Year.

Year	Sample A		Sample B		Sample C		Sample D		Sample E		Sample F		Sample G		Sample H		Sample I		Sample J		Sample K		Sample L1		Sample L2		Sample L3		Sample M1		Sample M2		Sample M3/4		
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	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	1,708	68.8					2,074	2,500											
2011	2,417	88.9	491	79.6	1,538	88.1	300	88.7	589	92.5	3,307	87.2	794	88.9	988	86.8			3,136				2,056	80.1	2,228	87.9	924								
2012	2,285	89.0	439	78.8	1,465	89.6	286	87.8	115	80.0	3,073	87.9	772	89.0	927	88.2			3,178	80.4	1,526		1,857	79.0	2,221	85.9	931	87.2							
2013	2,172	89.7	389	82.5	1,413	88.5	267	86.9	98	83.7	2,872	89.4	730	92.7	873	89.7			2,855	80.7	1,549	82.7	1,740	78.3	2,151	83.9	912	82.9	2,723						
2014	2,064	90.8	356	84.8	1,346	90.0	249	85.5	90	86.7	2,732	88.4	720	89.0	823	88.9			2,497	84.5	1,438	82.5	1,501	83.1	1,990	71.2	824	72.7	2,787	72.2					
2015	1,986	86.6	329	81.5	1,294	87.4	229	84.3	83	84.3	2,577	88.2	694	87.3	788	86.8			2,296	86.4	1,297	85.4	1,396	84.8	1,849	74.6	772	76.3	2,404	69.3	1,096				
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	1,201	87.1	1,280	87.7	1,712	73.9	719	75.0	2,062	72.4	1,054	62.6	3,320		

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

## **5.2 Predicting the Probability of Re-Interviewing vs. Refusal in the Year 2016**

Based on the household and interview characteristics measured in the year 2015, and some regional information measured in 2016, we aim at predicting the probability of agreement vs. refusal to participate in the survey for households that were re-contacted in 2016. 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 5.2. Table 6 reports the subsample-specific estimates of logit models for the probability of participating relative to refusing to participate. Note again that the estimates of regression models of the previous waves 1985 through 2015 are not reported in the present documentation due to space restrictions. These can as well be found in previous attrition reports.

Table 5.2: Definition of the Regressors of the Logit Model of Refusal

Variable	Label	Value
<b>Interview Characteristics</b>		
New Household	Household new in SOEP	0/1
New SOEP Member	Head of household had less than 4 interviews	0/1
Not Original Sample Memb.	Head of household is not an original sample member	0/1
New Address	Household moved	0/1
Subsample F2	Household from F2 ISOEP, Additional Foreigner Enumeration	0/1
Part. Unit Nonresponse	Household member(s) did not participate last wave	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
Successful Int. Related H.	Successful interview of related household	0/1
Interviewer Related H.	Same interviewer in related household	0/1
Change of Interviewer	Change of interviewer between the two last waves	0/1
Phone Unknown	Telephone number undisclosed	0/1
Email Known	Email address disclosed	0/1
Late Interview	Interview done in later months	0/1
Short Interview	Duration of the interview less than 10 Minutes	0/1
CAPI	Computer Assisted Personal Interview	0/1
Mother-Child-Questionnaire	Additional mother-child-questionnaire in household	0/1
DJ-Test	Completed DJ cognitive test	0/1
High Item Nonresp. H.	High item nonresponse in household questionnaire	0/1
Item Nonresponse Finan. Q.	Item Nonresponse on one or more financial questions	0/1
<b>Demographic Characteristics</b>		
Single Household	One-person household	0/1
Age 35-44	Head of household aged between 35 and 44 years	0/1
Age 55-64	Head of household aged between 55 and 64 years	0/1
Age 65-74	Head of household aged between 65 and 74 years	0/1
Partner in Household	Head of household lives with partner	0/1
Single	Head of household is single	0/1
<b>Work and Education</b>		
Evening Work	Household member(s) work often in the evening	0/1
Same employer 1st Q.	Head of household: duration working for the current employer 1st Q.	0/1
Same employer 2nd Q.	Head of household: duration working for the current employer 2nd Q.	0/1

Table 5.2 – Continued from previous page

Variable	Label	Value
Same employer 4th Q.	Head of household: duration working for the current employer 4th Q.	0/1
Job Intention	Head of household intends to obtain employment in the future	0/1
University Degree	Head of household has university degree	0/1
<b>Financial Situation and Insurance</b>		
Low Income	Low household income within the 1st quartile	0/1
Low Disposable Income	Low disposable income within the 1st quartile	0/1
High Disposable Income	High disposable income within the 3rd quartile	0/1
No Investments	The household did not own any investment securities last year	0/1
Investment Securities	The household owned 4 or more types of inv. securities last year	0/1
Private Health Insurance	Head of household is privately insured	0/1
Insurance Change	Head of household changed health insurance provider in prev. year	0/1
Donation	Head of household donated money last year	0/1
<b>Health</b>		
Health Restrictions	Household member(s) restricted in daily life due to health problems	0/1
Sick Note	Household member(s) was certified sick for more than 6 weeks last year	0/1
High Blood Pressure	Household member(s) diagnosed with high blood pressure	0/1
Chronic Back Pain	Household member(s) diagnosed with chronic back pain	0/1
Joint Disease	Household member(s) diagnosed with joint disease	0/1
Dementia	Household member(s) diagnosed with dementia	0/1
Person in need of care	Person in need of care lives in the household	0/1
<b>Personality, Attitudes and Other Characteristics</b>		
Worried	Head of household felt often worried in the last 4 weeks	0/1
Angry	Head of household felt often angry in the last 4 weeks	0/1
No Friends	Head of household does not have close friends	0/1
Many Friends	Head of household has more than ten close friends	0/1
Visited by Foreigner	Head of household was visited by foreigner(s) in prev. year	0/1
Strong Polit. Interest	Head of household is very interested in politics	0/1
Party Preference	Head of household has a preference for particular political party	0/1
Churchgoing	Head of Household attendet often religous events/church	0/1
No Religion	Head of household is not a member of a church or religious community	0/1
Peace Concerns	Head of household worried about the maintaining of peace	0/1
No Health Concerns	Head of household not worried about the own health	0/1
No Retirement Concerns	Household member(s) not worried about the own retirement pension	0/1
Antiforeignism Concerns	Household member(s) very worried about hostility to foreigners	0/1
No Antiforeignism Concerns	Household member(s) not worried about hostility to foreigners	0/1

Table 5.2 – Continued from previous page

Variable	Label	Value
Forgiving	Household member(s) tries to forgive and forget after being wronged	0/1
Revenge	Household member(s) will pay back if put in a difficult situation	0/1
No Revenge	Household member(s) will not pay back if put in a difficult situation	0/1
Quick recovery	Household member(s) gets over quickly after hurt feelings	0/1
No Revenge for Injustice	Household member(s) would not take a revenge if suffered a great wrong	0/1
Positive Self-Attitude	Household member(s) have positive attitude toward themselves	0/1
Self-Determination	Household member(s) thinks that how life goes depends on oneself	0/1
Success through Work	Household member(s) thinks that one has to work hard to succeed	0/1
Skills Importance	Household member(s) considers inborn skills more important than efforts	0/1
Efforts Importance	Household member(s) disagree that inborn skills are more imp. than efforts	0/1
Social Impact	Household member(s) disagree that political activism has social impact	0/1
Association Member	Household member(s) particip. in organisation/association	0/1
Trade Union Member	Household member(s) is member of trade union	0/1
Citizens' initiative	Household member(s) particip. in citizens' initiative/ polit. party	0/1
No Visit to Germans	Household member(s) did not visit german people in prev. year	0/1
<b>Building, Area, and Region</b>		
Subtenant	Household members are subtenants of the dwelling	0/1
Small Dwelling	Area of the dwelling is within the first tertile	0/1
Green Power	Household is buying electricity at a rate for green power	0/1
Academics Area	Household located in area with high share of academics	0/1
Low Share Abitur-Graduates	Household located in area with low share of abitur graduates	0/1
High Purchasing Power	Household located in area with high purchasing power	0/1
Single Households Area	Household located in area with single households	0/1
Family Households Area	Household located in area with high share of family households	0/1
Single Family Houses	Household located in area with single family houses	0/1
Eastern Europeans	Household located in area with high share of east. europeans/late repatriates	0/1
Islamic Countries	Household located in area with high share of people from isl. countr.	0/1
High Voter Turnout	Household located in area with high turnout during Federal election 2013	0/1
Low Voter Turnout	Household located in area with low turnout during Federal election 2013	0/1
SPD Voters Area	Household located in area with high share of SPD voters	0/1
Urban Area	Household located in a city with more than 100,000 inhabitants	0/1
Lower Saxony/Bremen	Household located in Lower Saxony or Bremen	0/1
Hessen	Household located in Hessen	0/1
Rhineland-Pal./Saarland	Household located in Rhineland-Palatinate or Saarland	0/1

Table 5.3: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2016

	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
Intercept	1,01***	-0,24	0,49***	0,38	1,30***	0,75***	1,83***	0,23	0,84***	0,92***	0,94***	0,68***	0,72***	0,34**	0,18*
<b>Interview Characteristics</b>															
New Household	-0,96***		-0,70*			-0,70***	-1,39***		-0,68**	-0,67*	-0,87*	-0,53*	-1,55***		
New SOEP Member			-0,84*			-0,62*						-0,42*			
Not Original Sample Memb.	-0,16*								-0,35**	-0,52**	-0,47***	-0,37***	-0,61***	-0,49***	
New Address							-0,56*		-0,29*						
Subsample F2						-0,51*									
Part. Unit Nonresponse						-0,31***				-0,39***	-0,30*	-0,34***			-0,48***
Temporary Drop-Out	-1,93***					-2,22***	-1,42***	-1,35***	-1,27***	-1,09***	-0,91***	-0,73***	-0,93***	-0,75***	
Drop Out Related Household			-0,45**				-1,01***								
Successful Int. Related H.												0,22*	0,41*		
Interviewer Related H.	0,29***	0,51*	0,58***												
Change of Interviewer	-0,34**	-0,69*				-0,44***	-0,72**		-0,26**	-0,38***	-0,44***	-0,57***			
Phone Unknown	-0,33**	-0,78**	-0,68***				-0,48*		-0,35*	-0,53**		-0,81***	-0,75***	-0,82***	-0,40***
Email Known														0,22**	
Late Interview	-0,28***				-0,97**										-0,21*
Short Interview													0,30*		
CAPI								0,47***				0,24***			
Mother-Child-Questionnaire														0,19*	
DJ-Test	-0,61*														
High Item Nonresp. H.	-0,20**												-0,49***		
Item Nonresponse Finan. Q.															-0,52**
<b>Demographic Characteristics</b>															
Single Household						-0,13*	-0,93***				-0,62**				
Age 35-44				-0,84*											
Age 55-64									0,27**						
Age 65-74									0,24**						
Partner in Household							-0,80**								
Single												-0,16*			
Evening Work			0,24*												

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



Table 5.3 – 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
Same employer 1st Q.			0,37**			0,15*									
Same employer 2nd Q.														0,25**	
Same employer 4th Q.						-0,26*									
Job Intention	-0,29*														
University Degree												0,22*			
<b>Financial Sit. and Insurance</b>															
Low Income								0,57***							
Low Disposable Income										0,50***					
High Disposable Income				1,01*											
No Investments								-0,64***							
Investment Securities									0,33*						
Private Health Insurance				-1,51**											
Insurance Change										-0,49*					
Donation						0,24***									
<b>Health</b>															
Health Restrictions														0,19**	
Sick Note														0,33*	
High Blood Pressure								0,54***							
Chronic Back Pain			0,39***												
Joint Disease			0,33**												
Dementia						-0,87**									
Person in need of care										-0,59**					
<b>Personality and Attitudes</b>															
Worried						-0,26*									
Angry		0,58*													
No Friends								-0,69***							
Many Friends														0,26*	
Visited by Foreigner														-0,20*	
Strong Polit. Interest					-0,69*										
Party Preference								0,36**			0,20*				
Churchgoing			-0,45**												
No Religion									0,24***						
Peace Concerns		0,45*													
No Health Concerns		-0,97***												0,19**	

Note: \*\*\*p &lt;0.01; \*\*p &lt;0.05; \*p &lt;0.10.

Table 5.3 – 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
No Retirement Concerns		0,73***													
Antiforeignism Concerns	0,19**														
No Antiforeignism Concerns												0,22*			
Forgiving														-0,15*	
Revenge														0,29**	
No Revenge														0,22**	
Quick recovery														0,15*	
No Revenge for Injustice						0,16**									
Positive Self-Attitude							0,37*								
Self-Determination			0,20*	0,66*											
Success through Work		0,58**													
Skills Importance											0,18*				
Efforts Importance															0,31*
Social Impact															0,23*
Association Member											0,19*				
Trade Union Member								0,43*							
Citizens' initiative															-0,34*
No Visit to Germans															-0,22*
<b>Building, Area, and Region</b>															
Subtenant	-0,44*					-0,38*									
Small Dwelling		0,58**													
Green Power									0,25*			0,25**			
Academics Area	0,22**										0,22*				
Low Share Abitur-Graduates		-0,47*													
High Purchasing Power								0,31*							
Single Households Area											0,22*				
Family Households Area								-0,50**							
Single Family Houses								0,32*							
Eastern Europeans														0,53**	
Islamic Countries		-0,40*													
High Voter Turnout												-0,22**			
Low Voter Turnout														0,29*	
SPD Voters Area															0,24*
Urban Area						0,22***									

Note: \*\*\*p &lt;0.01; \*\*p &lt;0.05; \*p &lt;0.10.

Table 5.3 – 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
Lower Saxony/Bremen		-0,91**													
Hessen														-0,35**	
Rhineland-Pal./Saarland						0,34*						-0,47**			0,44*
<i>No. of Observations</i>	1853	295	1213	124	80	2398	660	717	2108	1201	1280	1712	719	2067	1064
<i>Log Likelihood</i>	-569,46	-115,25	-359,25	-41,84	-26,62	-726,80	-146,41	-178,24	-579,70	-347,44	-391,03	-770,81	-303,40	-988,32	-659,89

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

## 6 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 6.1) and whether they are used in a given wave and subsample (Table 6.2). Tables 6.3 and 6.4 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 6.1: Marginal Distributions - Household Level

Variables	Marginal Distributions
Federal State <sup>21</sup> ( <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
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

Table 6.1: Marginal Distributions - Household Level

Variables	Marginal Distributions
Migration Second Generation ( <i>Migr.</i> )	no 2 <sup>nd</sup> generation migrant in household at least one 2 <sup>nd</sup> gen. migrant in household born after 1995 at least one 2 <sup>nd</sup> gen. migrant in household born 1975-1994 at least one 2 <sup>nd</sup> 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 <sup>22</sup> — the rest
Target Population AB ( <i>AB</i> )	Household size and country of origin ( <i>altogether 47 combinations</i> )
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 <sup>23</sup> 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
Children Typology ( <i>Child</i> )	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 ( <i>L1</i> )	<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)

Table 6.1: Marginal Distributions - Household Level

Variables	Marginal Distributions
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
Target Population H, J, K (H, J, K)	West Germany (without Berlin), all household members German West Germany, at least one household member without Ger. Nat. East Germany (incl. Berlin)

Table 6.1: Marginal Distributions - Household Level

Variables	Marginal Distributions
Household Size and Number of Employed Household Members <sup>24</sup> ( <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 <sup>25</sup>
	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
Origin and Year of Immigration ( <i>MI</i> ) <sup>26</sup>	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	

Table 6.1: Marginal Distributions - Household Level

Variables	Marginal Distributions
Origin and Year of Immigration (M2) <sup>26</sup>	2009-2011, Germany
	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

<sup>21</sup>Different 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

<sup>22</sup>An additional category "2010-2013" is used from 2015 on

<sup>23</sup>Deutschmark (DM)

<sup>24</sup>*Sample J*: sorted by East and West Germany

<sup>25</sup>Arbeitslosengeld II

<sup>26</sup>Personal 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 6.2: Margins - Household Level

Samples	Fed. State	Mun. Size	H. Size	Owner	H. Type	Migr.	Nat.	Imm. Year	AB	E, F	G
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							
2007 A-H	+	+	+	+							
2008 A-H	+	+	+	+							
2009 A-I	+	+	+	+							
2010 A-L2	+ * L1 L2	+ * L1 L2	+ *	+ *	+ *	+ *					
2011 A-L3	+ * L3 J	+ * L3 J	+ * J	+ * J	+ * J	+ * J					
2012 A-K	+ * K	+ * K	+ * K	+ * K	* K	+ * K					
2013 A-M1	+ * M1	+ * M1	+ * M1	+ *	+ *	+ *	+	+			
2014 A-M1	+	+	+	+	+	+	+	+			
2015 A-M2	+ * M2	+ * M2	+ * M2	+ *	+ *	+ *	+ *	+ *			
2016 A-M3/4	+ *	+ *	+ *	+ *	+ *	+ *	+ *	+ *			

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

Table 6.2: Margins - Household Level

Samples	Child	L1	L2	L1/L2	L3	L1/L3	H, J, K	Empl.	ALG	Reg.	M1	M2
1984 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												
1999 A-E												
2000 A-F												
2001 A-F												
2002 A-G												
2003 A-G												
2004 A-G												
2005 A-G												
2006 A-H							* H					
2007 A-H												
2008 A-H												
2009 A-I												
2010 A-L2	L1 L2	* L1	* L2	L1 L2								
2011 A-L3	L3				* L3	* L3	* J	J	J			
2012 A-K							* K	K	K			
2013 A-M1										M1	M1	
2014 A-M1												
2015 A-M2										M2		M2
2016 A-M3/4												

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

Table 6.3: 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 <sup>27</sup> 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 2<sup>nd</sup> 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 nat. and born before 1975
Age <sup>28</sup>	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+

Table 6.3: Marginal Distributions - Person Level

Variables	Distributions
Gender	Male — Female
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 — 2010-2013<sup>29</sup> — the rest</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>

Table 6.3: Marginal Distributions - Person Level

Variables	Distributions
Gender, Origin and Year of Immigration (M1)	1st generation, earlier than 1995, Turkey, m/f <sup>30</sup>
	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

Table 6.3: Marginal Distributions - Person Level

Variables	Distributions
Gender, Origin and Year of Immigration (M2)	did not immigrate between 2009-2013, m/f
	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	
Federal State (Fed. State)	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	
Date of arrival in Germany (Arrival)	2013 Q1 — 2013 Q2 — 2013 Q3 — 2013 Q4
	2014 Q1 — 2014 Q2 — 2014 Q3 — 2014 Q4
	2015 Q1 — 2015 Q2 — 2015 Q3 — 2015 Q4

Table 6.3: Marginal Distributions - Person Level

Variables	Distributions
Country of origin (M3/4)	Syria Afghanistan Iraq Albania, Serbia, Kosovo Eritrea, Somalia Iran, Pakistan Other
Refugee (Ref.)	Came to Germany as refugee between 2013 and 2016 (yes/no)

<sup>27</sup>Deutschmark (DM)

<sup>28</sup>Different categorisation:

*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)

*Sample M3/4*: 3 categories between 15 and 24: 15-17, 18-19, 21-24. Last category is 60+

<sup>29</sup>An additional category "2010-2013" is used from 2015 on

<sup>30</sup>Each category distinguishes between male (m) or female (f) gender of the respondent

Table 6.4: Margins - Person Level

	Sample	Age and Gender	H. Type	German	G	Migrant 2 <sup>nd</sup> Gen.	Age	Gender	L1	L2	L3	Nation.	Imm. Year	Migrant	M1	M2	Fed. State	Arrival	M3/4	Ref.
<b>1984</b>	A-B	+	+	+																
<b>1985</b>	A-B	+	+	+																
<b>1986</b>	A-B	+	+	+																
<b>1987</b>	A-B	+	+	+																
<b>1988</b>	A-B	+	+	+																
<b>1989</b>	A-B	+	+	+																
<b>1990</b>	A-B	+	+	+																
<b>1991</b>	A-B	+	+	+																
<b>1992</b>	A-B	+	+	+																
<b>1993</b>	A-B	+	+	+																
<b>1994</b>	A-B	+	+	+																
<b>1995</b>	A-B	+	+	+																
<b>1996</b>	A-B	+	+	+																
<b>1997</b>	A-B	+	+	+																
<b>1998</b>	A-E	+ * E	+ * E	+ * E																
<b>1999</b>	A-E	+	+	+																
<b>2000</b>	A-F	+ * F	+ * F	+ * F																
<b>2001</b>	A-F	+	+	+																
<b>2002</b>	A-G	+ *	+ *	+ *																* G
<b>2003</b>	A-G	+	+	+																
<b>2004</b>	A-G	+	+	+																
<b>2005</b>	A-G	+	+	+																
<b>2006</b>	A-H	+ * H	+ * H	+ * H																

Note: (+) margins for standard weights; (\*) margins for standard weights without the new samples;

(*sample letter*) margins for standalone weights of a new sample



Table 6.4: Margins - Person Level

	Sample	Age and Gender	H. Type	German	G	Migrant 2 <sup>nd</sup> Gen.	Age	Gender	L1	L2	L3	Nation.	Imm. Year	Migrant	M1	M2	Fed. State	Arrival	M3/4	Ref.
<b>2007</b>	A-H	+	+	+																
<b>2008</b>	A-H	+	+	+																
<b>2009</b>	A-I	+	+	+																
<b>2010</b>	A-L2	+ *	+ *	+ *		+ *	L1 L2	L1 L2	L1	L2										
<b>2011</b>	A-L3	+ * J	+ * J	+ * J		+ * J	L3	L3			L3									
<b>2012</b>	A-K	+ * K	+ * K	+ * K		+ * K														
<b>2013</b>	A-M1	+ *	+ *	+ *		+ *	M1					+ *	+ *	M1	* M1					
<b>2014</b>	A-M1	+	+	+		+						+	+							
<b>2015</b>	A-M2	+ *	+ *	+ *		+ *	M2					+ *	+ *			M2				
<b>2016</b>	A-M3/4	+ *	+ * M3/4	+ *		+ *	M3/4	M3/4				+ *	+ *				M3/4	M3/4	M3/4	+

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

## 7 Summary Statistics of the Derived Longitudinal and Cross-Sectional Weights

Based on the regression models of successful vs. unsuccessful recontacts and agreements vs. refusals to participate, we derive two sets of predicted probabilities, the product of which is the household's "staying probability". The inverse of the probability of staying in the SOEP in 2016 based on characteristics measured in 2015, variable `BGHBLEIB`, lends itself as a longitudinal weighting variable which itself corrects for selective attrition between waves 2015 and 2016. Tables 7.1, Table 7.2, Table 7.3, Table 7.4 and Table 7.5 report some subsample specific summary statistics of the longitudinal weights in each wave.

The product of the cross-sectional weight in 2015, variable `BFHHRF`, and the longitudinal weight in 2016, variable `BGHBLEIB`, provide the raw data for the cross-sectional weight in 2016. 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 2016.

Tables 7.6 and 7.7 report subsample specific summary statistics of the derived cross-sectional weighting variable `BGHHRF` and in comparison all previous cross-sectional weights `AHHRF` through `BFHHRF`.

Table 7.1: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples A through D (Percentiles of \$HBLEIB up to Wave BG).

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

Table 7.2: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples E through G (Percentiles of \$HBLEIB up to Wave BG).

Year	Sample E				Sample F				Sample G			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
1999	1.00	1.23	1.47	886								
2000	1.03	1.07	1.21	838								
2001	1.01	1.05	1.25	811	1.08	1.14	1.59	4,911				
2002	1.01	1.02	1.20	773	1.03	1.05	1.46	4,586				
2003	1.04	1.04	1.15	744	1.02	1.04	1.24	4,386	1.06	1.10	1.17	911
2004	1.00	1.01	1.08	732	1.02	1.03	1.19	4,235	1.02	1.03	1.25	904
2005	1.01	1.03	1.18	706	1.01	1.03	1.17	4,070	1.03	1.06	1.25	879
2006	1.00	1.03	1.21	686	1.01	1.03	1.29	3,895	1.00	1.04	1.31	859
2007	1.01	1.01	1.16	647	1.01	1.03	1.15	3,694	1.01	1.05	1.17	824
2008	1.00	1.01	1.19	602	1.01	1.03	1.14	3,513	1.01	1.03	1.18	787
2009	1.00	1.04	1.17	574	1.02	1.04	1.34	3,303	1.02	1.04	1.36	757
2010	1.01	1.04	1.25	553	1.01	1.05	1.40	3,055	1.00	1.01	1.23	743
2011	1.00	1.00	1.17	545	1.01	1.05	1.34	2,885	1.00	1.03	1.35	706
2012	1.05	1.24	1.66	92	1.02	1.08	1.30	2,702	1.02	1.07	1.24	687
2013	1.07	1.20	1.32	82	1.01	1.06	1.21	2,567	1.02	1.05	1.15	677
2014	1.03	1.03	1.42	78	1.02	1.05	1.25	2,414	1.01	1.07	1.32	641
2015	1.13	1.13	1.42	70	1.01	1.05	1.30	2,273	1.01	1.07	1.38	606
2016	1.06	1.06	1.38	68	1.03	1.08	1.24	2,094	1.02	1.02	1.26	590

Table 7.3: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples H, J and K (Percentiles of \$HBLEIB up to Wave BG).

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

Table 7.4: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples L1, L2 and L3 (Percentiles of \$HBLEIB up to Wave BG).

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	1.03	1.12	1.45	1,247	1.09	1.25	1.69	1,416	1.10	1.24	1.80	599
2015	1.01	1.06	1.36	1,184	1.04	1.15	1.92	1,379	1.03	1.12	1.77	589
2016	1.02	1.08	1.25	1,122	1.04	1.16	1.97	1,265	1.03	1.15	1.68	539

Table 7.5: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples M1 and M2 (Percentiles of \$HBLEIB up to Wave BG).

Year	Sample M1				Sample M2			
	p10	p50	p90	N	p10	p50	p90	N
2014	1.03	1.11	1.59	1,362				
2015	1.03	1.12	1.45	1,247				
2016	1.01	1.06	1.36	1,184	1.04	1.15	1.92	1,379

Table 7.6: Summary Statistics of the Derived Cross-Sectional Weights at the Household Level (Percentiles of \$HHRF up to Wave 33).

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	682	3,917	5,081	6,430	8,471	10,033	5,322
1986	541	752	3,598	5,302	6,838	9,280	11,117	5,090
1987	546	790	3,543	5,381	7,044	9,575	11,455	5,026
1988	534	804	3,560	5,638	7,541	10,355	12,539	4,814
1989	549	819	3,597	5,840	7,879	10,810	13,274	4,690
1990	699	1,073	2,217	4,600	7,042	9,897	12,393	6,819
1991	680	1,043	2,333	4,693	7,145	10,289	12,875	6,699
1992	667	1,028	2,335	4,660	7,134	10,529	13,651	6,665
1993	688	1,056	2,401	4,670	7,257	10,750	13,979	6,637
1994	709	1,102	2,400	4,674	7,284	11,208	14,719	6,559
1995	696	1,110	2,387	4,359	6,977	11,087	14,843	6,768
1996	732	1,165	2,387	4,360	7,010	11,384	15,318	6,699
1997	737	1,210	2,399	4,317	7,050	11,848	15,874	6,621
1998	980	1,351	2,331	3,974	6,224	9,891	13,119	7,492
1999	967	1,322	2,309	3,984	6,497	10,884	14,344	7,220
2000	800	1,101	1,758	2,525	3,568	5,084	6,519	13,082
2001	752	1,030	1,754	2,752	4,145	6,095	7,822	11,796
2002	507	657	1,221	2,554	4,192	6,513	8,257	12,320
2003	505	673	1,236	2,562	4,325	6,827	9,084	11,909
2004	491	669	1,214	2,536	4,420	7,266	9,829	11,644
2005	490	678	1,234	2,544	4,523	7,586	10,884	11,294
2006	458	649	1,267	2,388	4,139	6,885	9,766	12,361
2007	457	651	1,254	2,467	4,470	7,582	10,685	11,552
2008	459	656	1,275	2,554	4,754	8,226	11,565	10,921
2009	472	666	1,301	2,628	5,036	9,072	12,424	10,270
2010	220	360	666	1,429	3,646	7,366	11,084	13,888
2011	214	325	611	1,502	3,091	5,605	7,816	16,703
2012	218	329	636	1,635	3,160	5,712	7,634	16,397
2013	170	266	514	1,283	2,904	5,382	7,464	17,992
2014	196	309	614	1,507	3,337	6,240	8,732	15,946
2015	183	294	610	1,471	3,360	6,530	9,259	15,908
2016	37	74	329	1,145	3,088	6,160	9,348	17,762

Table 7.7: Summary Statistics of the Derived Cross-Sectional Weights at the Person Level (Percentiles of \$PHRF up to Wave 33).

Year	p5	p10	p25	p50	p75	p90	p95	N
1984	397	553	1,174	4,364	5,224	6,052	6,800	16,173
1985	453	634	1,421	4,621	5,720	6,891	8,062	14,508
1986	488	678	1,532	4,683	6,024	7,587	9,000	13,804
1987	509	718	1,597	4,726	6,223	7,894	9,406	13,563
1988	488	686	1,623	4,894	6,564	8,504	10,199	12,872
1989	528	754	1,747	5,019	6,893	8,969	10,691	12,443
1990	681	1,023	1,900	3,441	6,147	8,273	10,233	18,254
1991	731	1,072	1,915	3,708	6,193	8,475	10,599	17,844
1992	777	1,139	2,000	3,741	6,300	8,723	11,111	17,429
1993	845	1,232	2,088	3,835	6,385	9,010	11,396	17,072
1994	871	1,286	2,109	3,846	6,416	9,272	12,067	16,715
1995	765	1,145	2,016	3,612	6,097	9,068	12,163	17,345
1996	797	1,190	2,025	3,642	6,145	9,421	12,769	16,944
1997	839	1,213	2,062	3,667	6,249	9,698	13,347	16,583
1998	914	1,267	2,041	3,512	5,591	8,515	11,405	18,249
1999	902	1,240	2,015	3,492	5,795	9,246	12,664	17,501
2000	723	974	1,565	2,314	3,216	4,561	5,853	30,784
2001	689	935	1,535	2,451	3,647	5,409	6,926	27,956
2002	445	612	1,064	2,191	3,714	5,819	7,628	29,101
2003	439	620	1,086	2,205	3,814	6,124	8,225	27,867
2004	436	616	1,085	2,185	3,893	6,512	8,854	26,918
2005	440	630	1,119	2,237	4,014	6,876	9,599	25,638
2006	413	596	1,115	2,176	3,678	6,303	8,826	27,442
2007	413	599	1,122	2,228	3,921	6,951	10,099	25,505
2008	424	611	1,153	2,298	4,126	7,648	11,249	23,792
2009	440	627	1,183	2,386	4,358	8,367	12,573	22,096
2010	175	278	535	1,034	2,519	5,405	8,581	35,945
2011	165	252	457	986	2,390	4,421	6,622	42,031
2012	168	251	471	1,100	2,542	4,512	6,747	40,351
2013	138	210	405	886	2,227	4,291	6,350	44,633
2014	153	240	477	1,077	2,572	4,920	7,354	38,839
2015	140	224	466	1,083	2,570	5,067	7,661	38,224
2016	22	36	177	786	2,276	4,812	7,384	44,151

## 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.
- Haisken-DeNew, J. and J. Frick (2017). *Desktop Companion to the Socio Economic Panel (SOEP)*. Web version under construction: <http://about.paneldata.org/soep/dtc/>.
- 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.



- 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 Auffrischungstichproben am Beispiel des Sozio-oekonomischen Panels (SOEP)*. In: *Nonresponse Bias: Qualitätssicherung sozialwissenschaftlicher Umfragen*. Ed. by J. Schupp and C. Wolf. Wiesbaden: Springer. pp.409–444.
- 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, and R. Siegers (2017a). *Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2015)*. SOEP Survey Papers 408, DIW/SOEP, Berlin 2017.
- Kroh, M., S. Kühne, J. Jacobsen, M. Siegert, and R. Siegers (2017b). *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.
- Pffor, 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. Trüding, and B. Rammstedt (2015). “Are Incentive Effects on Response Rates and Non-response 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 Ergänzungstichprobe 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.