960 2021

SOEP Survey Papers
Series C - Data Documentations (Datendokumentationen)

SOEP-Core v36 – Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2019)

Rainer Siegers, Hans Walter Steinhauer, Lennart Dührsen

DIW SOEP

Running since 1984, the German Socio-Economic Panel (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**

Please cite this paper as follows:

Rainer Siegers, Hans Walter Steinhauer, Lennart Dührsen. 2021. SOEP-Core v36 – Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2019). SOEP Survey Papers 960: Series C. Berlin: DIW/SOEP

ISSN: 2193-5580 (online)

DIW Berlin
German Socio-Economic Panel (SOEP)
Mohrenstr. 58
10117 Berlin
Germany

soeppapers@diw.de

Data Documentation:

Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2019)

Rainer Siegers

Hans Walter Steinhauer

Lennart Dührsen

April 16, 2021

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

Contents

List of Tables

List of Figures

Introduction

This data documentation is meant to provide SOEP users with a general overview of the longitudinal development of the survey over the past 36 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 person 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, the Refugee Samples M3/4 and M5, and Samples N and O. Because of their short-running time series, the latest Samples (P, Q) are not outlined separately, but will be added over the next years. For a general overview on the integration of enlargement and refreshment samples into the SOEP see Kroh et al. (2015b).

The SOEP study surveys not only the original sample from the first wave, but also households and persons that entered the survey at later points in time. They enter, for example, when SOEP households split (i.e., individuals move out and form their own households), when people move into SOEP households, and when an original sample member gives birth to a "new sample member". For a detailed review of the SOEP inclusion rules for new sample units and their treatment within the weighting framework see Spiess et al. (2008) and Schonlau et al. (2011).

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

Households may leave the survey for several reasons. SOEP's weighting strategy distinguishes between survey-related reasons and reasons unrelated to the survey (for a detailed description of the SOEP weighting strategy, see Rendtel (1995) and Schonlau et al. (2013) and for a general overview, Kara et al. (2018)). We ignore panel attrition of the latter form due to respondents moving abroad or dying, since these cases technically represent an exit from the underlying population. The third section of this paper provides initial evidence on the risk of survey-related panel attrition in different groups of the original sample units (e.g., in different subsamples, age, educational, and income groups).

The fourth section reports in more detail on the occurrence of unsuccessful follow-ups to household addresses by cross-section and subsample, and subsample-specific regression models of the probability of unsuccessful follow-ups in 2019 based on the characteristics of households measured in 2018. 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, Siegers et al. (2020) for wave BI).

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 2019: BJHBLEIB and BJPBLEIB. Based on the inverse probability of observing households and persons in 2018, the staying probability in 2019, and additional post-stratification to meet benchmarks of known margins of the underlying population in 2019, we derive the cross-sectional weights BJHHRF and BJPHRF.

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

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

1 Sampling of SOEP Subsamples A to Q

1.1 Sample A (1984)

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

Key Facts

Sampling Design	Multistage stratified sampling procedure based on the ADM-Design[1]		
first stage	Stratification: federal states (NUTS 1)		
	governmental regions (NUTS 2)		
	municipality size		
	Clustering: 585 primary sampling units (PSUs)		
second stage	Random walk in each PSU		
	Selected unit: household		
Sample Size[2]		households	persons (thereof children)
	NET	4,524	11,366 *(2,290)*
	GROSS	7,430	
Field Period	February to October 1984		
Initial Survey Mode	Paper-and-Pencil Interviewing (PAPI), possibility for self-completion		
Number of Interviewers	592		
Initial Response Rate[3]	60.9%		
Initial Weighting Factor	Average	SD	min / max
	5,491	1,402	3,358 / 11,041
Further Readings	Infratest Sozialforschung (2011). *SOEP 1984 – Methodenbericht zum Befragungsjahr 1984 (Welle 1) des Sozio-oekonomischen Panels.* SOEP Survey Papers 1, DIW/SOEP, Berlin 2011.		

[1] ADM is the "Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V." (Working Group of the German Market and Social Research Institutes). For more information, see https://www.adm-ev.de/leistungen/arbeitsgemeinschaft-adm-stichproben/

[2] 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).

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

1.2 Sample B (1984)

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

Key Facts

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

1.3 Sample C (1990)

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

Key Facts

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

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

1.4 Sample D (1994/95)

Sample D "Immigrants" covers private households in which at least one household member had moved from abroad to West Germany after 1984. This sample includes two subsamples that were drawn independently in 1994 (D1) and in 1995 (D2).
The fieldwork organization sampled a small number of households of Sample D (*N=98*) drawing on a respondent-driven sampling procedure. In these 98 cases, inclusion probabilities cannot be derived directly and we thus do not assign weights to these households.

Key Facts

Sampling Design	Households with at least one person who moved to Germany since 1984 were identified in representative face-to-face and telephone surveys of the German population driven by Infratest and following the ADM-Design.			
Sample D1 (1994)	All eligible households which agreed to be re-contacted by the SOEP-Survey were selected for the gross sample. The gross sample was supplemented with 98 additional cases, which were obtained by a respondent-driven procedure.			
Sample D2 (1995)	Here a distinction was made between ethnic German immigrants from Eastern Europe as well as the GDR and *Other Immigrants*. While in case of Other Immigrants again all eligible households, that agreed to be re-contacted by the SOEP-Survey, were selected for the gross sample, among ethnic German immigrants approx. 70% were selected in order to compensate for overrepresentation of the latter subpopulations in Sample D1.			
Sample Size		households	persons (thereof children)	
		D1 D2	D1	D2
	NET	236 295[5]	719 *(248)*	905 *(283)*
	GROSS	307 385		
Field Period	January to March 1994 (D1) and January to April 1995 (D2)			
Initial Survey Mode	Paper-and-Pencil Interviewing (PAPI), possibility for self-completion			
Number of Interviewers	83 (1994)	206 (1995)		
Initial Response Rate	76.9% (D1)	76.6% (D2)		
Initial Weighting Factor	Average	SD	min / max	
(in 1995)	3,906	1,717	1,699 / 9,855	

[5] 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.

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.

1.5 Sample E (1998)

Sample E "Refreshment I" is the first sample that was designed to be representative for all private households in both East and West Germany. It is the first of several regular refreshment samples drawn to increase the overall size of the SOEP, compensate for panel-attrition and cover population changes, e.g. due to migration.
It is also the first sample in which the Computer-Assisted Personal Interviewing (CAPI) was implemented. Interviews in Samples A-D at this time were completely conducted using Paper-and-Pencil Interviewing (PAPI). To study mode effects, households of sample E were randomly allocated to CAPI and PAPI mode.
With the data distribution of 2012, parts of sample E have been extracted into the SOEP Innovation Sample.

Key Facts

Sampling Design	Multistage stratified sampling procedure based on the ADM-Design		
first stage	Stratification: federal states (NUTS 1)		
	governmental regions (NUTS 2)		
	municipality size		
	Clustering: 125 PSUs		
second stage	Random walk in each PSU		
	Selected unit: household		
Sample Size		households	persons (thereof children)
	NET	1,056	2,376 *(466)*
	GROSS	1,969	
Field Period	April to September 1998		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI) and Paper-and-Pencil Interviewing (PAPI)		
Number of Interviewers	130		
Initial Response Rate	53.6%		
Initial Weighting Factor	Average	SD	min / max
	35,568	18,294	14,827 / 205,099

Further Readings

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

Projektgruppe Das Sozio-oekonomische Panel (DIW) (1998). *Funktion und Design einer Ergänzungsstichprobe für das Sozio-oekonomische 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.

1.6 Sample F (2000)

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

Key Facts

Sampling Design	Multistage stratified sampling procedure based on the ADM-Design		
first stage	Stratification: federal states (NUTS 1)		
	governmental regions (NUTS 2)		
	counties (NUTS 3)		
	municipality size		
	Clustering: 985 PSUs		
second stage	Random walk in each PSU		
	Oversampling of “non-German" households		
	Selected unit: household		
Sample Size		households	persons (thereof children)
	NET	6,043	13,871 *(2,991)*
	GROSS	11,862	
Field Period	March to October 2000		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI) and Paper-and-Pencil Interviewing (PAPI), possibility for self-completion		
Number of Interviewers	671		
Initial Response Rate	50.9%		
Initial Weighting Factor	Average	SD	min / max
	6,364	2,224	2,376 / 18,861
Further Readings			

Infratest Sozialforschung (2011). *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.

1.7 Sample G (2002)

The 2002 Sample G "High Income" covers private households in Germany with a monthly income of at least DM[6]7,500 (EUR 3,835), which - due to the lack of an adequate sampling frame - were identified using a telephone screening procedure. From Wave 2 in 2003 onwards, only households with a net monthly income of at least EUR 4,500 were interviewed further.

Key Facts

Sampling Design	Households with a monthly income of ≥ DM 7,500 were identified in representative face-to-face and telephone surveys of the German population driven by Infratest and following the ADM-Design.		
first stage	From all 5,663 eligible households 3,672 were drawn, stratified by income and region (east/west) with oversampling of higher incomes and regions in East-Germany. Of these 2,495 households agreed to be re-contacted by the SOEP-Survey and became the gross sample.		
Sample Size		households	persons (thereof children)
	NET	1,224	3,364 *(693)*
	GROSS	2,493	
Field Period	March to July 2002		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI) and Paper-and-Pencil Interviewing (PAPI), possibility for self-completion		
Number of Interviewers	276		
Initial Response Rate	49.1%		
Initial Weighting Factor	Average	SD	min / max
	2,084	953	983 / 9,757
Further Readings			

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

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

[7] Deutschmark (DM)

1.8 Sample H (2006)

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

Key Facts

Sampling Design	Multistage stratified sampling procedure based on the ADM-Design		
first stage	Stratification: federal states (NUTS 1)		
	governmental regions (NUTS 2)		
	municipality size		
	Clustering: 250 PSUs		
second stage	Random walk in each PSU		
	Selected unit: household		
Sample Size		households	persons (thereof children)
	NET	1,506	3,239 *(623)*
	GROSS	3,747	
Field Period	March to July 2006		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	243		
Initial Response Rate	40.2%		
Initial Weighting Factor	Average	SD	min / max
	26,443	13,453	9,024 / 12,8852
Further Readings			

TNS Infratest Sozialforschung (2011). *SOEP 2006 – Methodenbericht Erstbefragung der Ergänzungsstichprobe H zum Befragungsjahr 2006 (Welle 23) des Sozio-oekonomischen Panels.* SOEP Survey Papers 57, DIW/SOEP, Berlin 2011.

1.9 Sample I (2009)

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

Key Facts

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

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

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

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

1.10 Sample J (2011)

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

Key Facts

Sampling Design	Multistage stratified sampling procedure based on the ADM-Design		
first stage	Stratification: federal states (NUTS 1)		
	governmental regions (NUTS 2)		
	municipality size		
	Clustering: 307 PSUs		
second stage	Random walk for address listing in each PSU		
	Oversampling of migrant households[8] such that the share of migrants for each PSU is doubled		
	Selected unit: household		
Sample Size		households	persons (thereof children)
	NET	3,136	6,308 *(1,147)*
	GROSS	9,492	
Field Period	March to October 2011		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	338		
Initial Response Rate	33.0%		
Initial Weighting Factor	Average	SD	min / max
	12,592	6,166	1,950 / 49,307
Further Readings			

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

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

[8] Identification of potentially migrant households using onomastic procedure.

1.11 Sample K (2012)

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

Key Facts

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

Further Readings

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

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

1.12 Sample L1 (FiD) (2010)

Sample L1 "Cohort Sample"[9], covers private households in Germany, in which at least one household member is a child that was born between January 2007 and March 2010. Again, migrants identified by an "onomastic procedure" are oversampled.

Key Facts

Sampling Design	Multistage stratified sampling procedure based on information from local registration offices (Einwohnermeldeämter)		
first stage	Stratification: federal states (NUTS 1) governmental regions (NUTS 2) municipality size		
	Clustering: 159 PSUs		
second stage	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[10] such that the share of migrants for each PSU is doubled		
	Selected unit: child in the respective cohort		
Sample Size		households	persons (thereof children)
	NET	2,074	7,670 *(3,900)*
	GROSS	5,286	
Field Period	June to October 2010		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	204		
Initial Response Rate	39.2%		
Initial Weighting Factor	Average	SD	min / max
	935	576	75 / 3,494
Further Readings			

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

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

[9] 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.

[10] Identification of potentially migrant addresses using onomastic procedure and information on the citizenship.

1.13 Sample L2 (FiD) (2010)

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

Key Facts

Sampling Design	Persons in potentially eligible households were identified in representative face-to-face and telephone surveys of the German population following the ADM-Design. Telephone screening (CATI-Screening) was then conducted in order to verify the eligibility and willingness of the households to participate. Selected unit: person		
Sample Size		households	persons (thereof children)
	NET	2,500[11]	8,838 *(4,611)*
	GROSS	3,281	
Field Period	March to June 2010		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	343		
Initial Response Rate	76.2%		
Initial Weighting Factor	Average	SD	min / max
	1,596	1,035	213 / 7,702
Further Readings			

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

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

[11] 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.

1.14 Sample L3 (FiD) (2011)

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

Key Facts

Sampling Design

Persons in potentially eligible households were identified in representative face-to-face and telephone surveys of the German population following the ADM-Design. Telephone screening (CATI-Screening) was then conducted to verify the eligibility and willingness of the households to participate.
Selected unit: person

Sample Size

	households	persons (thereof children)
NET	924[12]	3,579 *(2,092)*
GROSS	1,144	

Field Period

March to June 2011

Initial Survey Mode

Computer-Assisted Personal Interviewing (CAPI)

Number of Interviewers

250

Initial Response Rate

80.8%

Initial Weighting Factor

Average	SD	min / max
2,359	1,582	468 / 12,146

Further Readings

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

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

[12] 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.

1.15 Sample M1 (2013)

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

Key Facts

Sampling Design	Multistage stratified sampling design based on the IEB database		
first stage	Stratification: federal states (NUTS 1)		
	county type (urban/rural)		
	Clustering: 250 PSUs proportional to number of migrants[13] in each stratum		
second stage	Simulated random walk algorithm in each PSU		
	Disproportional address sampling according to country of origin and migration generation		
	Selected unit: person		
Sample Size		households	persons (thereof children)
	NET	2,723	7,445 *(2,481)*
	GROSS	11,051	
Field Period	May to November 2013		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	232		
Initial Response Rate	35.0%[14]		
Initial Weighting Factor	Average	SD	min / max
	1,564	1,540	63 / 9,117

Further Readings

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

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

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

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

[13] Identification of target persons using information on nationality, BA measures and onomastic procedure.
[14] Including the 1,145 households that were screened out and not taken into further consideration.

1.16 Sample M2 (2015)

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

Key Facts

Sampling Design	Multistage stratified sampling design based on the IEB database		
first stage	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[15] in each stratum		
second stage	Disproportional address sampling in each PSU according to country of origin		
	Selected unit: person		
Sample Size		households	persons (thereof children)
	NET	1,096	2,638 *(927)*
	GROSS	6,008	
Field Period	May to December 2015		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	143		
Initial Response Rate	32.6%[16]		
Initial Weighting Factor	Average	SD	min / max
	901	741	54 / 3,760
Further Readings			
	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.		

[15] Identified by the year they entered the IEB and former and current citizenship.

[16] Including the 863 households that were screened out and not taken into further consideration.

1.17 Sample M3/4 (2016)

The 2016 "IAB-BAMF-SOEP Refugee Survey" (Samples M3 and M4) is a joint project of the *Institute for Employment Research* (IAB), the *Research Centre of the Federal Office for Migration and Refugees* (BAMF-FZ) as well as the SOEP. The target population of the samples consists of households with individuals who arrived in Germany between January 2013 and January 2016 and applied for asylum or were hosted as part of specific programs of the federal states (irrespective of their asylum procedure and their current legal status).
The first part of the sample (M3) was financed with funds from the research budget of the *Federal Employment Agency* (BA) allocated to the IAB. Sample M4 was funded by the *Federal Ministry of Education and Research* (BMBF) and has a focus on refugee families.

Key Facts

Sampling Design	Multistage stratified sampling design based on the German Central Register of Foreigners (AZR)		
first stage	Stratification: federal states (NUTS 1)		
	county type (urban/rural)		
	Clustering: 99 PSUs (M3) / 95 PSUs (M4)		
second stage	Disproportional address sampling in each PSU according to country of origin, current legal status, age and gender		
	Selected unit: person		
Sample Size		households	persons (thereof children)
	NET	3,273	9,856 *(5,391)*
	GROSS	6,761	
Field Period	June to December 2016		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	162		
Initial Response Rate	48.4%		
Initial Weighting Factor	Average	SD	min / max
	142	192	4 / 2,967

Further Readings

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

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

1.18 Sample M5 (2017)

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

Key Facts

Sampling Design	Multistage stratified sampling design based on the German Central Register of Foreigners (AZR)		
first stage	Stratification: federal states (NUTS 1)		
	county type (urban/rural)		
	Clustering: 99 PSUs		
second stage	Disproportional address sampling in each PSU according to country of origin, current legal status, gender, and target population (refreshment vs. enlargement)		
	Selected unit: person		
Sample Size		households	persons (thereof children)
	NET	1,519	4,161 *(1,909)*
	GROSS	2,871	
Field Period	June to October 2017		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	33		
Initial Response Rate	52.9%		
Initial Weighting Factor	Average	SD	min / max
	138	162	5/2189
Further Readings			
	Jacobsen, J., M. Kroh, S. Kühne, J. A. Scheible, R. Siegers, and M. Siegert (2019). *Supplementary of the IAB-BAMF-SOEP Survey of Refugees in Germany (M5) 2017.* SOEP Survey Papers 605, DIW/SOEP, Berlin 2019.		

1.19 Sample N (2017)

Participants of Sample N were initially drawn in the context of the international *Project in Assessment of Adult Skills and Competencies* (PIAAC) in 2012 that was initiated by the OECD[17]. The survey of the German subsample was carried out by the *Leibniz-Institute for the Social Sciences* (GESIS) and the target population of PIAAC 2012 Germany consisted of adults from age 16 through 65 that lived in Germany (on the reference date of 1 December 2011). The fieldwork in 2012 resulted in a net sample of 5,319 persons. Participants were then transferred into the PIAAC-L panel study[18], which followed the concept of "Anchor Persons", meaning that only original PIAAC sample members were followed in subsequent waves. The waves of PIAAC-L surveyed not only the PIAAC anchor persons, but other household members as well and already introduced items similar to those of the SOEP. The respective waves were conducted in the years 2014 (3,758 anchor interviews), 2015 (3,263) and 2016 (2,967), of which 2,811 anchor persons have agreed to be transferred into the SOEP. Finally, Sample N is based on respondents that took part in the last wave of PIAAC-L in 2016 and gave consent to be transferred into the SOEP.

Key Facts

Sampling Design[19]	Two-staged stratified and clustered sampling procedure based on information from local registration offices (Einwohnermeldeämter)		
first stage	Stratification: federal states administrative regions districts county type (rural/urban)		
	Clustering: 277 PSUs		
second stage	systematic random sampling in each PSU		
	Selected unit: person[20]		
Sample Size[21]		households	persons (thereof children)
	NET	2,378[22]	4,807 *(1,037)*
	GROSS	3,447	
Field Period	March to August 2017		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	287		
Initial Response Rate	69.0%		
Initial Weighting Factor	Average	SD	min / max
	14,016	11,455	1,755 / 157,875

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

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

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

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

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

[22] 64 of these households will be realised the first time in wave 2 of Sample N.

Further Readings

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

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

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

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

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

1.20 Sample O (2018)

Sample O is a refreshment sample that is aimed at evaluating the urban development and planning program 'Soziale Stadt'. The target population of Sample O consists of all households located in one of the 'Soziale Stadt' areas. The corresponding households have been sampled using spatially referenced data. Besides a novel sampling approach, the refreshment sample itself provides an additional data infrastructure for urban and regional planning and research.

Key Facts

Sampling Design	Shape files restricting residential areas in which households were sampled as well as information about number and coordinates of buildings within these areas have been provided by *the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR).*		
first stage	Stratification: 20 Regions (by Federal states and population size)		
	Clustering: PSUs		
second stage	Within the PSUs buildings were randomly selected.		
third stage	Within each of the selected buildings households were selected using the Kish selection grid.		
	Selected unit: household		
Sample Size		households	persons (thereof children)
	NET	935	1,730 *(479)*
	GROSS	6,119	
Field Period	March to August 2018		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	122		
Initial Response Rate	15.3%		
Initial Weighting Factor	Average	SD	min / max
	2,559	2,032	89 / 9440
Further Readings			
	Steinhauer, H. W., M. Kroh, and J. Goebel (2020). *SOEP-Core – 2018: Sampling, Nonresponse, and Weighting in Sample O.* SOEP Survey Papers 827: SOEP Survey Papers Series C –Data Documentation, DIW/SOEP, Berlin 2020.		

1.21 Sample P (2019)

Sample P "Top Shareholder Sample", covers households in Germany in which at least one household member belongs to the top percentile in terms of the estimated value of his or her cumulative company shareholdings.

Key Facts

Sampling Design	Multistage stratified sampling design based on the global company database ORBIS, which was provided by the business information publisher Bureau van Dijk (BvD).		
first stage	Stratification: 24 Regions (by Federal states and population density) Clustering: 250 PSUs		
second stage	Disproportional address sampling in each PSU according to age, sex and estimated value of shareholdings Selected unit: person		
Sample Size		households	persons (thereof children)
	NET	1,960	3,589 *(1,149)*
	GROSS	22,728	
Field Period	January 2019 to February 2020		
Initial Survey Mode	Computer-Assisted Personal Interviewing (CAPI)		
Number of Interviewers	259		
Initial Response Rate	8.6%		
Initial Weighting Factor	Average	SD	min / max
	322	291	18 / 2,548
Further Readings			

Schröder, C., C. Bartels, K. Göbler, M. M. Grabka, J. König, R. Siegers and S. Zinn (2020). *Improving the Coverage of the Top-Wealth Population in the Socio-Economic Panel (SOEP)*. SOEP Papers on Multidisciplinary Panel Data Research 1114, DIW/SOEP, Berlin 2020.

1.22 Sample Q (2019)

The 2019 boost sample Q supplemented the SOEP core sample by queer households, including gender and sexual minorities such as lesbian, gay, bisexual, and trans* respondents. To recruit these households, a random telephone screening of adults living in Germany was conducted. Sample Q was funded by the *Federal Ministry of Education and Research (BMBF)*.

Key Facts

Sampling Design

Persons eligible to the target population were identified through nationwide omnibus surveys conducted by Kantar Public. A dual-frame method was used which makes it possible to also include respondents who only have a cell phone but not a landline. Subsequently, a telephone screening (CATI screening) was conducted to verify the eligibility and willingness of the target respondents (and their households) to participate.

Selected unit: person

Sample Size

	households	persons (thereof children)
NET	477	636 *(70)*
GROSS	813	

Field Period

September 2018 to August 2019 (telephone screening)
April to November 2019 (interviews in households)

Initial Survey Mode

Computer-Assisted Personal Interviewing (CAPI)

Number of Interviewers

221

Initial Response Rate

58.7%

Initial Weighting Factor

Average	SD	min / max
3,157	2,192	582 / 12,734

Further Readings

De Vries, L., M. Fischer, M. Kroh, S. Kühne and D. Richter. (2021). *Design, Nonresponse, and Weighting in the 2019 Sample Q (Queer) of the Socio-Economic Panel*. SOEP Survey Papers 940: SOEP Survey Papers Series C – Data Documentation, DIW/SOEP, Berlin 2021.

2 Developments in Sample Size

With respect to developments in sample size, the following figures focus on (2.1) comparing the number of successful interviews by cross-section, (2.2) providing a longitudinal study of panel attrition among the original sample members, (2.3) showing the entrance of new sample members by birth / moving into SOEP households and their participation behavior, (2.4) reporting share of original households in relation to new households from splits and (2.5) assessing the risk of survey-related attrition of original sample respondents by social characteristics.

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

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

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

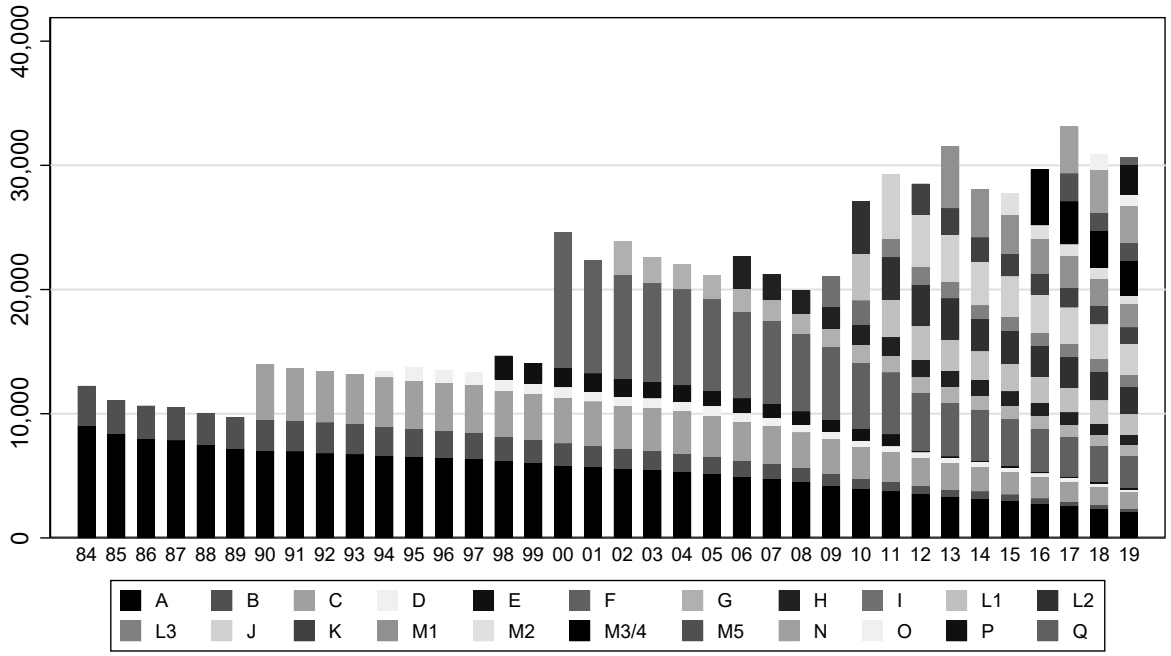

Figure 1: The Number of Successful Interviews with Persons by Subsamples A through Q, Waves 1 to 36.

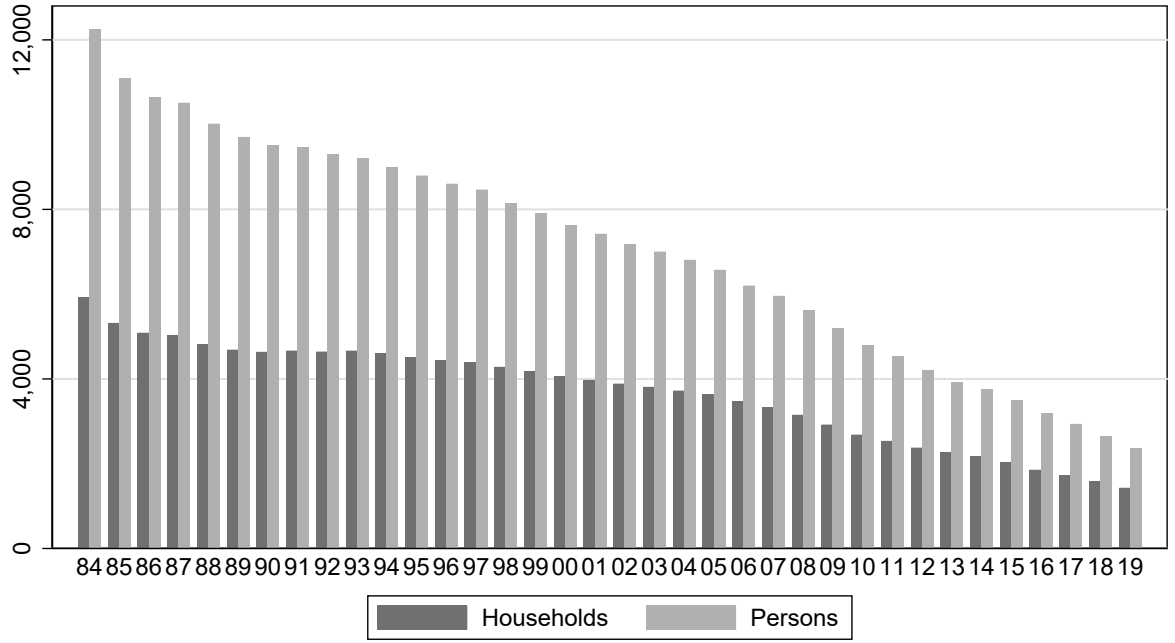

Figure 2: Comparison of Successful Interviews with Persons and Households (Subsamples A and B), Waves 1 to 36

Year	**1984**	**1985**	**1986**	**1987**	**1988**	**1989**	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**
Persons	12,245	11,090	10,646	10,516	10,023	9,710	9,519	9,467	9,305	9,206	9,001	8,798	8,606	8,467	8,145	7,909	7,623
Households	5,921	5,322	5,090	5,026	4,814	4,690	4,640	4,669	4,645	4,667	4,600	4,508	4,445	4,389	4,285	4,183	4,060

Year	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Persons	7,424	7,175	7,004	6,811	6,575	6,203	5,961	5,626	5,197	4,793	4,541	4,204	3,926	3,761	3,497	3,187	2,940	2,653	2,370
Households	3,977	3,889	3,814	3,724	3,635	3,476	3,337	3,154	2,923	2,686	2,539	2,379	2,270	2,176	2,028	1,857	1,729	1,581	1,433

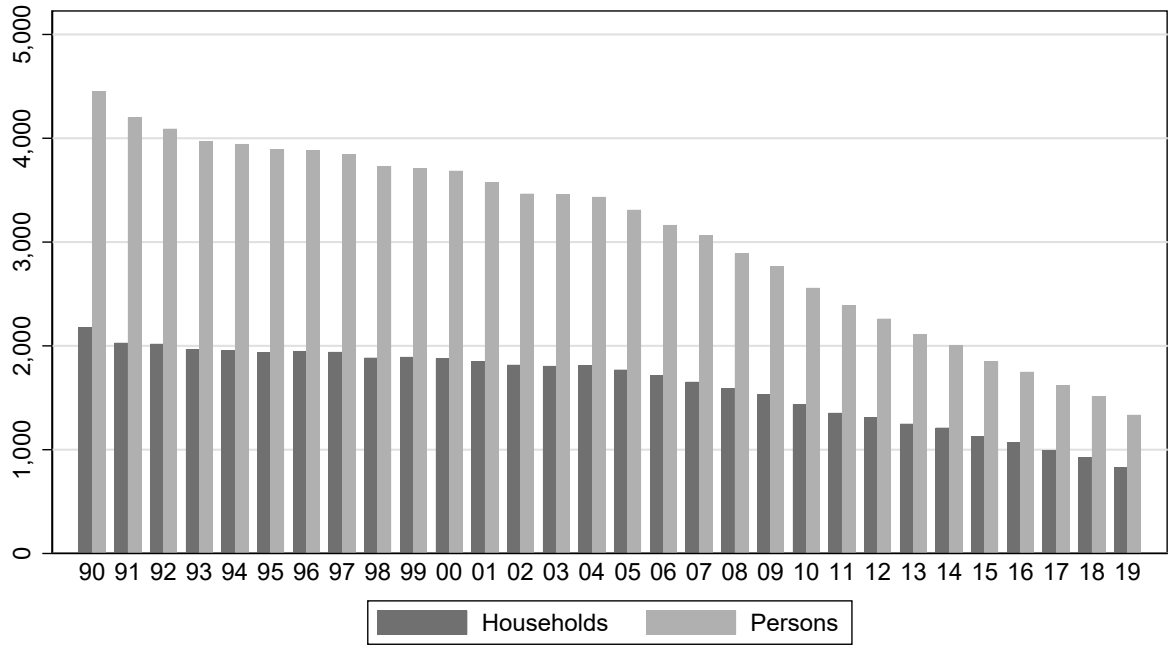

Figure 3: Comparison of Successful Interviews with Persons and Households (Subsample C), Waves 1 to 30

Year	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**
Persons	4,453	4,202	4,092	3,973	3,945	3,892	3,882	3,844	3,730	3,709	3,687	3,576	3,466	3,459
Households	2,179	2,030	2,020	1,970	1,959	1,938	1,951	1,942	1,886	1,894	1,879	1,850	1,818	1,807

Year	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Persons	3,435	3,311	3,165	3,067	2,892	2,769	2,559	2,392	2,262	2,111	2,006	1,853	1,750	1,622	1,516	1,336
Households	1,813	1,771	1,717	1,654	1,592	1,535	1,437	1,355	1,312	1,250	1,212	1,131	1,073	997	929	830

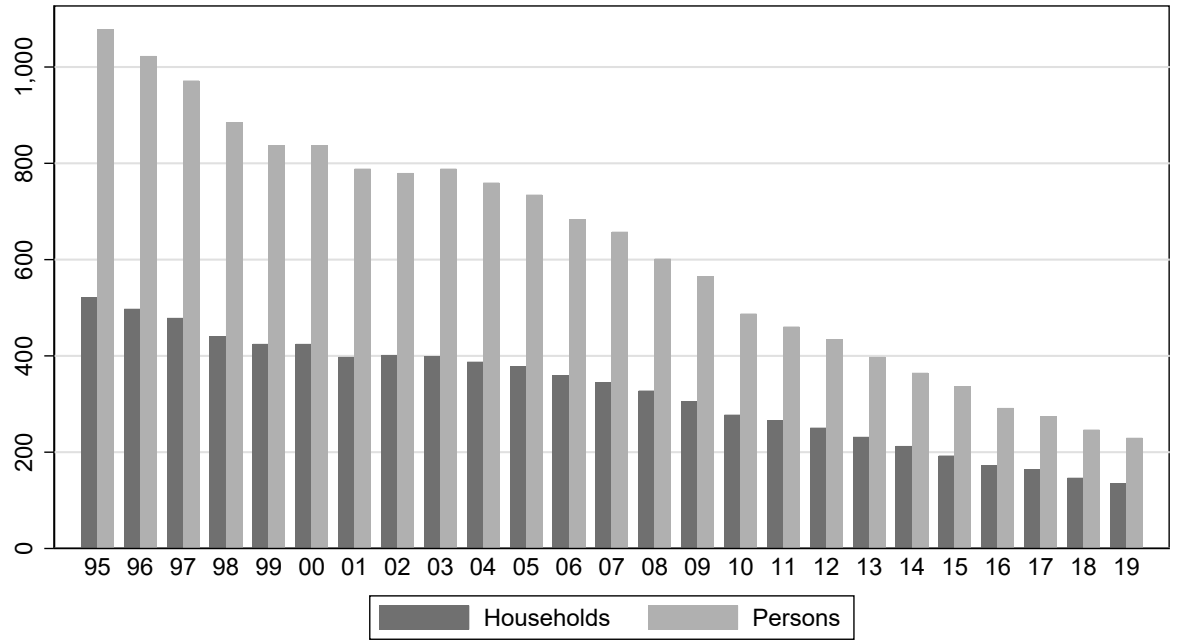

Figure 4: Comparison of Successful Interviews with Persons and Households (Subsample D), Waves 1 to 25

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

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Persons	658	602	565	488	461	435	398	365	337	292	275	247	230
Households	345	328	306	278	266	251	232	213	193	173	165	147	136

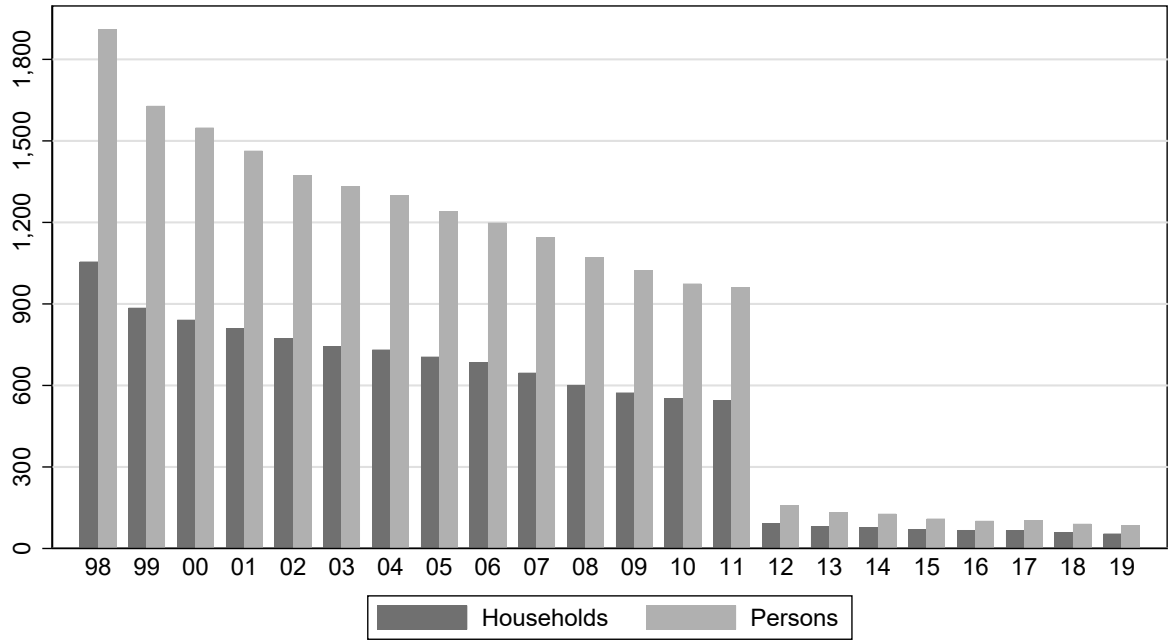

Figure 5: Comparison of Successful Interviews with Persons and Households (Subsample E), Waves 1 to 22[23]

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

Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Persons	1,071	1,024	975	961	160	134	128	110	102	104	91	85
Households	602	574	553	545	92	82	78	70	68	67	59	55

[23] 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 from 2012 onwards.

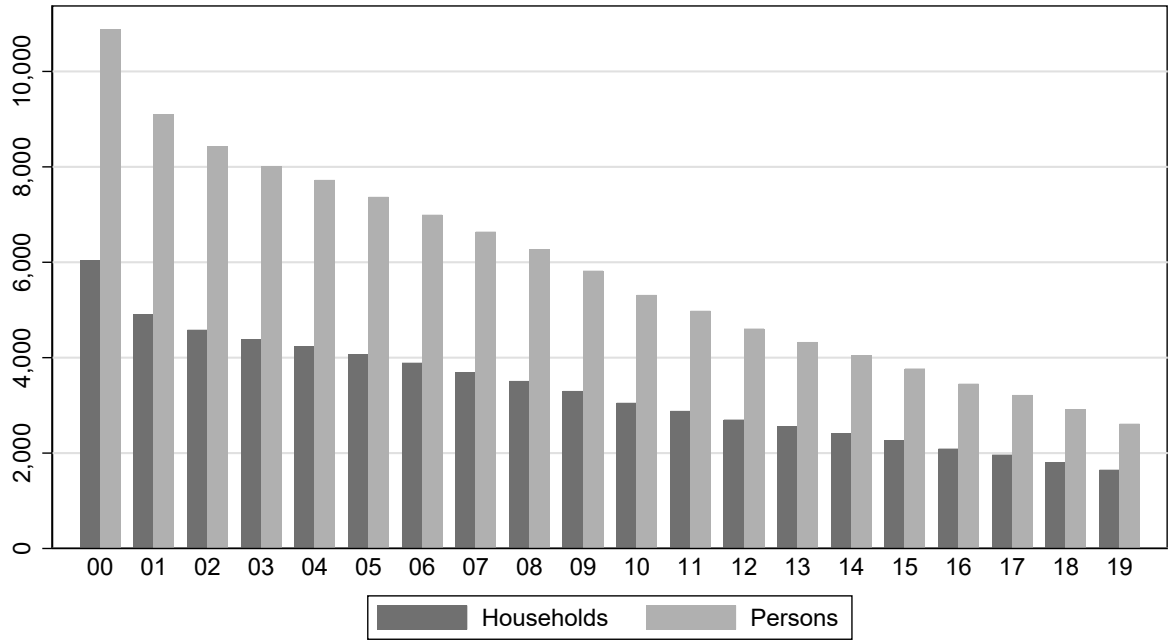

Figure 6: Comparison of Successful Interviews with Persons and Households (Subsample F), Waves 1 to 20

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Persons	10,880	9,098	8,427	8,010	7,727	7,372	6,997	6,642	6,276	5,824
Households	6,043	4,911	4,586	4,386	4,235	4,070	3,895	3,694	3,513	3,303

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Persons	5,316	4,984	4,610	4,329	4,049	3,773	3,455	3,219	2,923	2,616
Households	3,055	2,885	2,702	2,567	2,414	2,273	2,094	1,968	1,811	1,652

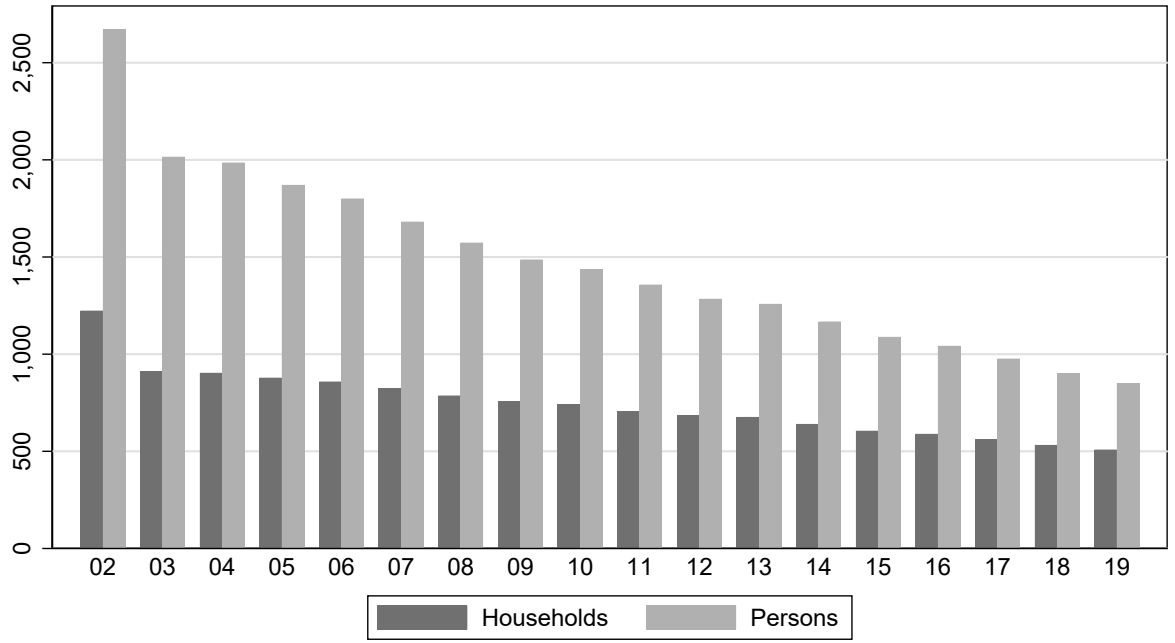

Figure 7: Comparison of Successful Interviews with Persons and Households (Subsample G), Waves 1 to 18[24]

Year	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Persons	2,671	2,016	1,986	1,871	1,801	1,682	1,574	1,487	1,438	1,358	1,285	1,259	1,168	1,089	1,043	977	903	851
Households	1,224	911	904	879	859	824	787	757	743	706	687	677	641	606	590	561	533	509

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

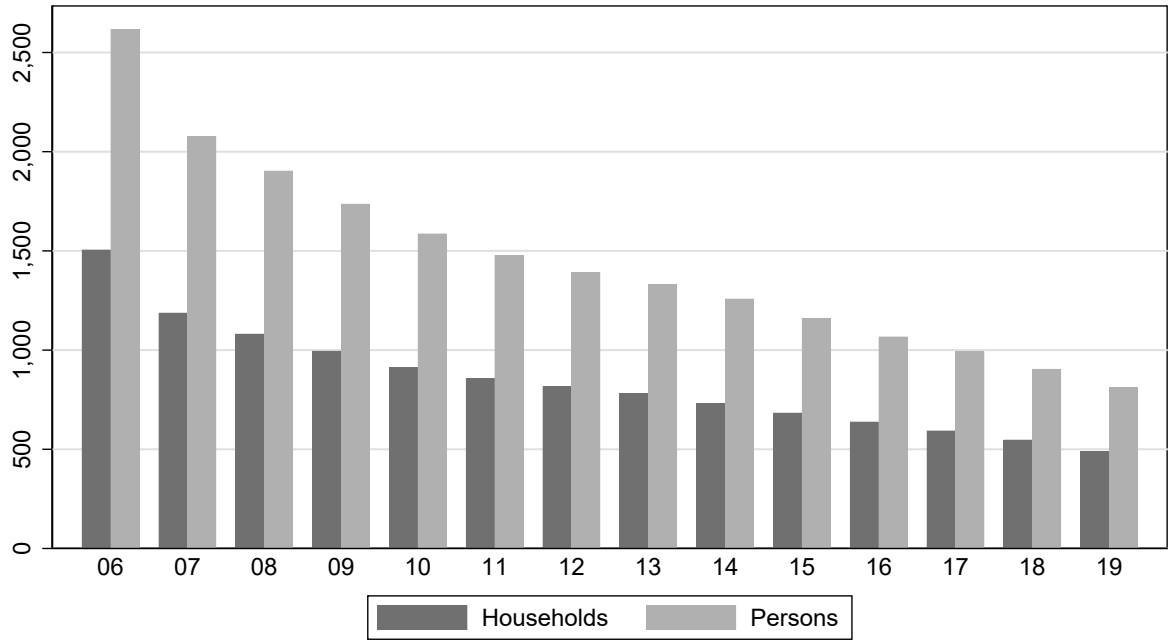

Figure 8: Comparison of Successful Interviews with Persons and Households (Subsample H), Waves 1 to 14

Year	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Persons	2,616	2,077	1,904	1,737	1,587	1,478	1,392	1,333	1,259	1,162	1,068	993	905	814
Households	1,506	1,188	1,082	996	913	858	818	783	732	684	639	594	548	491

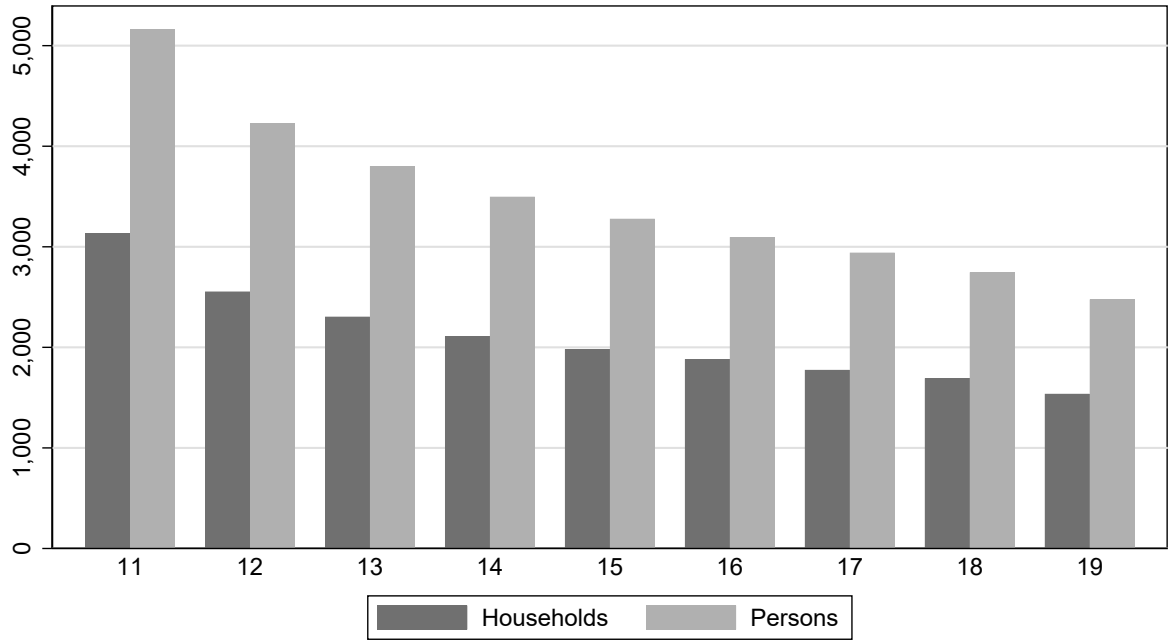

Figure 9: Comparison of Successful Interviews with Persons and Households (Subsample J), Waves 1 to 9

Year	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Persons	5,161	4,229	3,801	3,498	3,279	3,096	2,942	2,746	2,476
Households	3,136	2,555	2,305	2,110	1,983	1,883	1,776	1,692	1,538

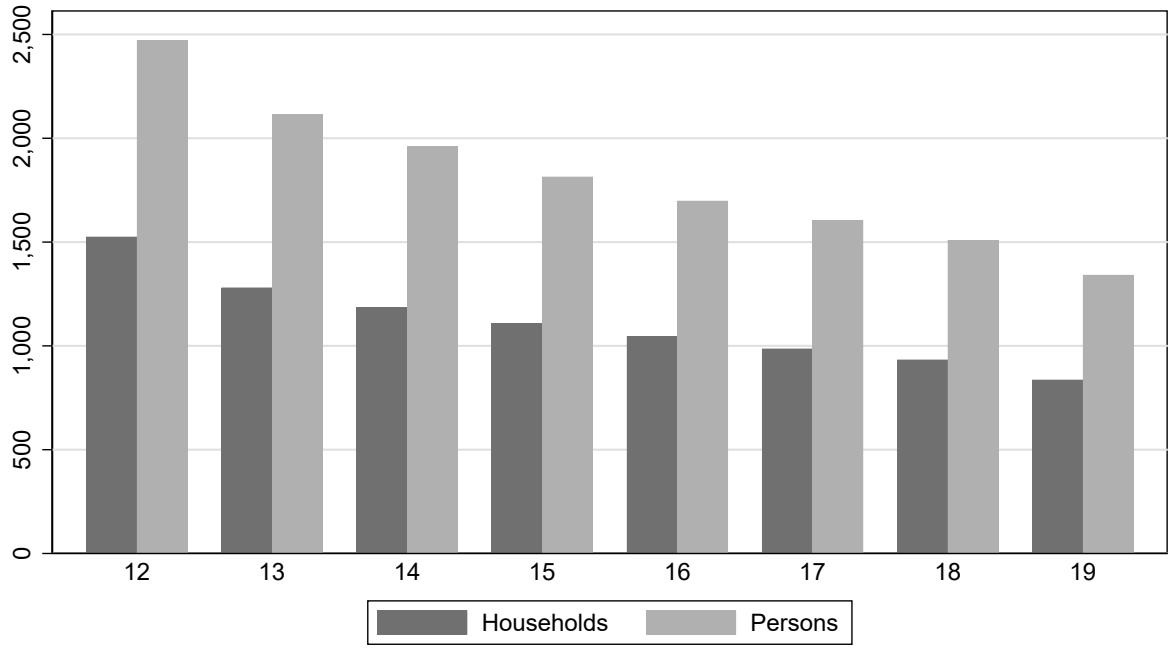

Figure 10: Comparison of Successful Interviews with Persons and Households (Subsample K), Waves 1 to 8

Year	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Persons	2,473	2,115	1,962	1,815	1,699	1,605	1,510	1,342
Households	1,526	1,281	1,187	1,108	1,046	987	934	837

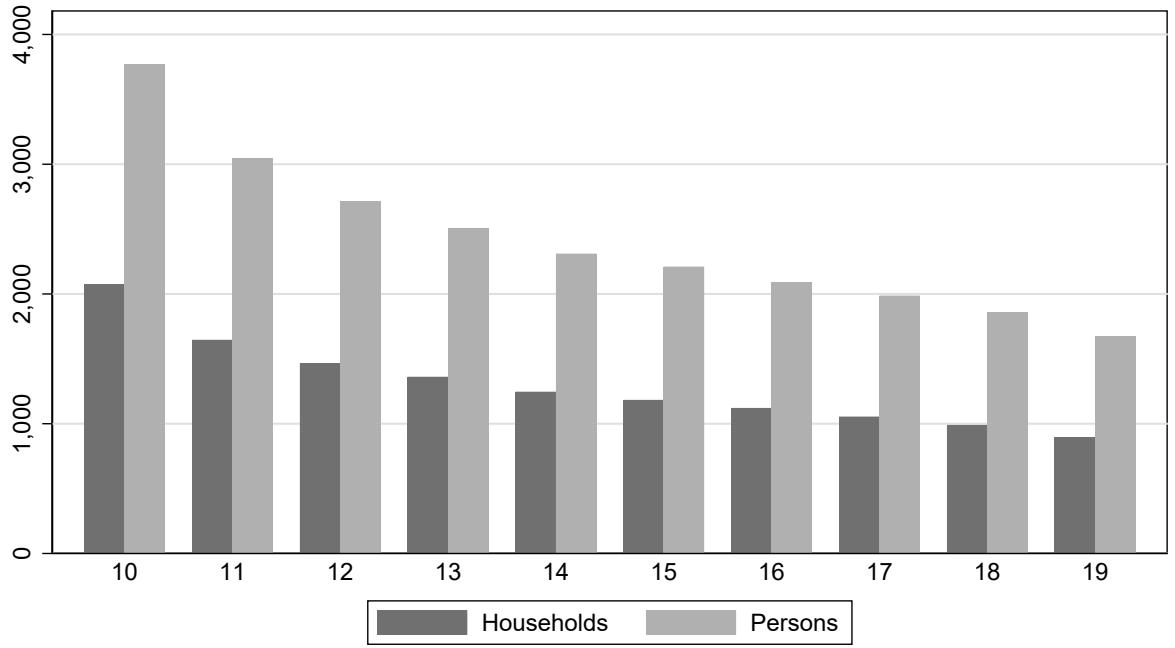

Figure 11: Comparison of Successful Interviews with Persons and Households (Subsample L1), Waves 1 to 10

Year	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Persons	3,770	3,048	2,713	2,506	2,311	2,211	2,091	1,988	1,861	1,675
Households	2,074	1,647	1,467	1,362	1,247	1,184	1,122	1,055	991	894

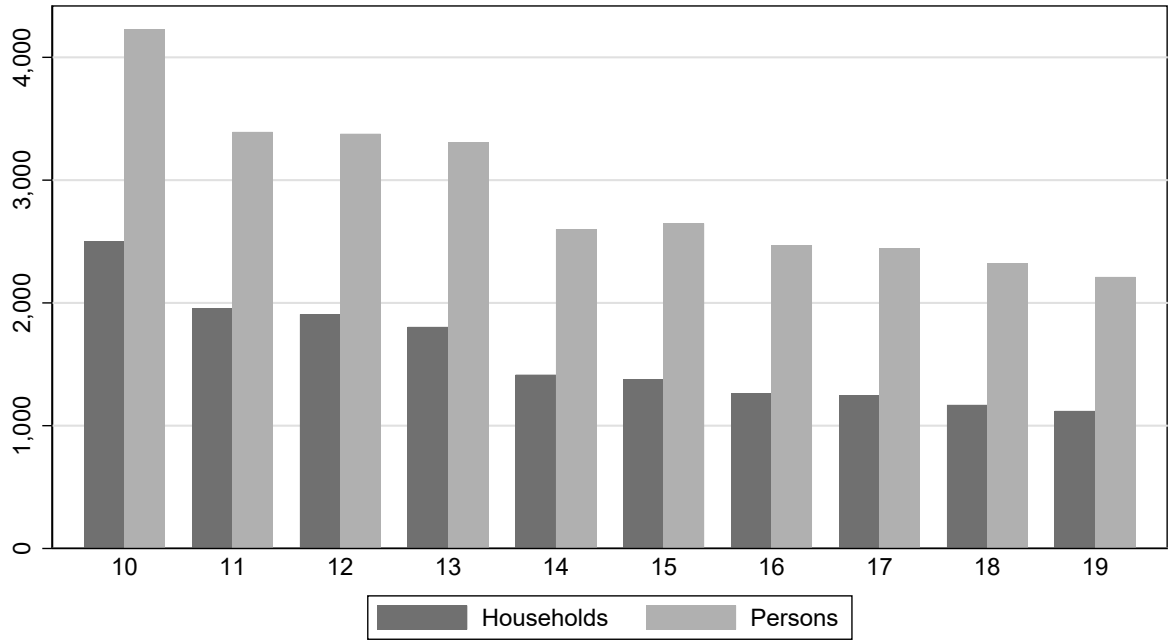

Figure 12: Comparison of Successful Interviews with Persons and Households (Subsample L2), Waves 1 to 10 [25,26]

Year	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Persons	4,227	3,393	3,378	3,307	2,600	2,647	2,469	2,447	2,324	2,212
Households	2,500	1,958	1,907	1,805	1,416	1,379	1,265	1,247	1,170	1,121

[25] 237 households were identified not to be part of the target population and were not followed in the second wave.
[26] In 2014 the default interview mode changed to Computer-Assisted Web Interviewing (CAWI).

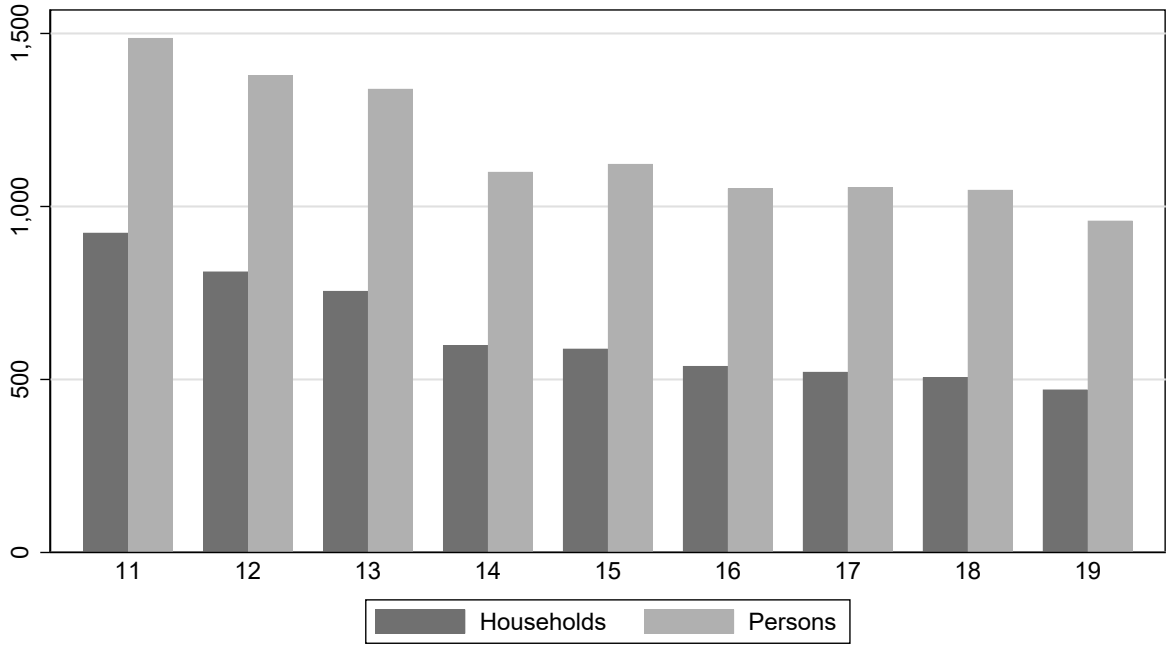

Figure 13: Comparison of Successful Interviews with Persons and Households (Subsample L3), Waves 1 to 9[27]

Year	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Persons	1,487	1,379	1,340	1,100	1,123	1,052	1,056	1,048	959
Households	924	812	756	599	589	539	522	506	471

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

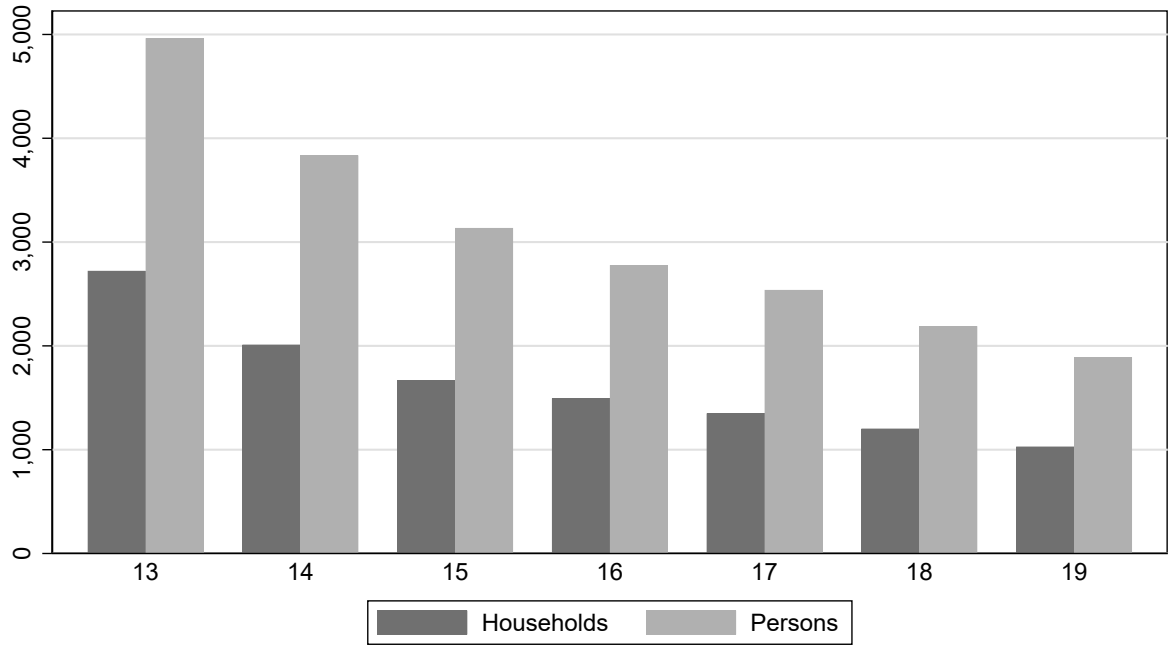

Figure 14: Comparison of Successful Interviews with Persons and Households (Subsample M1), Waves 1 to 7

Year	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Persons	4,964	3,835	3,136	2,778	2,539	2,190	1,891
Households	2,723	2,012	1,667	1,493	1,350	1,203	1,030

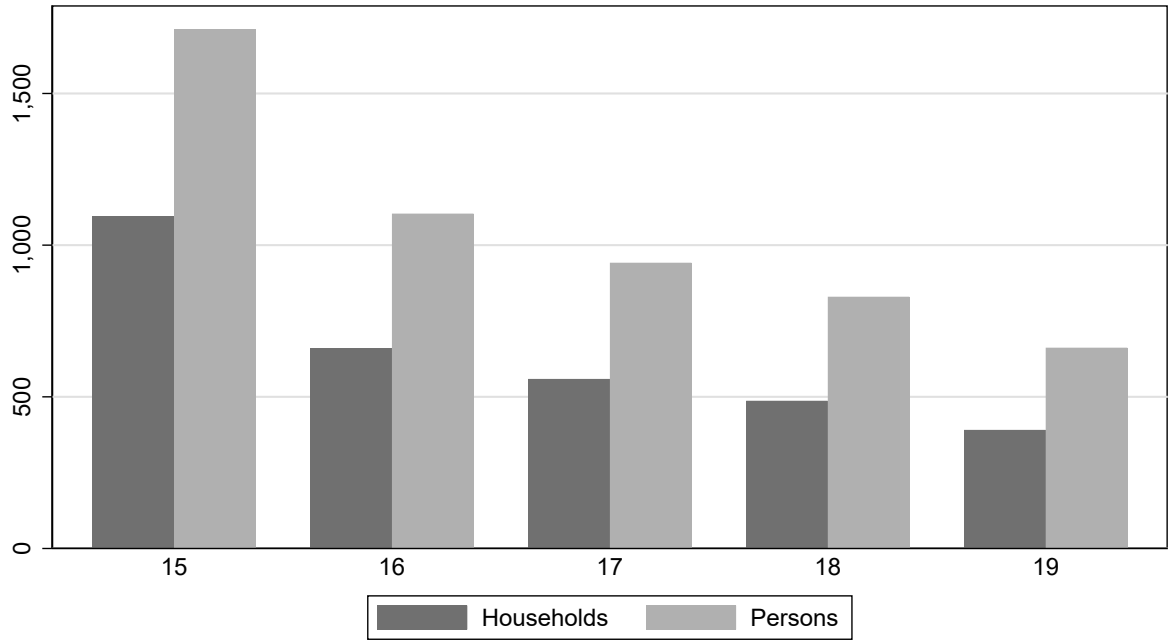

Figure 15: Comparison of Successful Interviews with Persons and Households (Subsample M2), Waves 1 to 5

Year	**2015**	**2016**	**2017**	**2018**	**2019**
Persons	1,711	1,104	942	830	662
Households	1,096	660	559	487	391

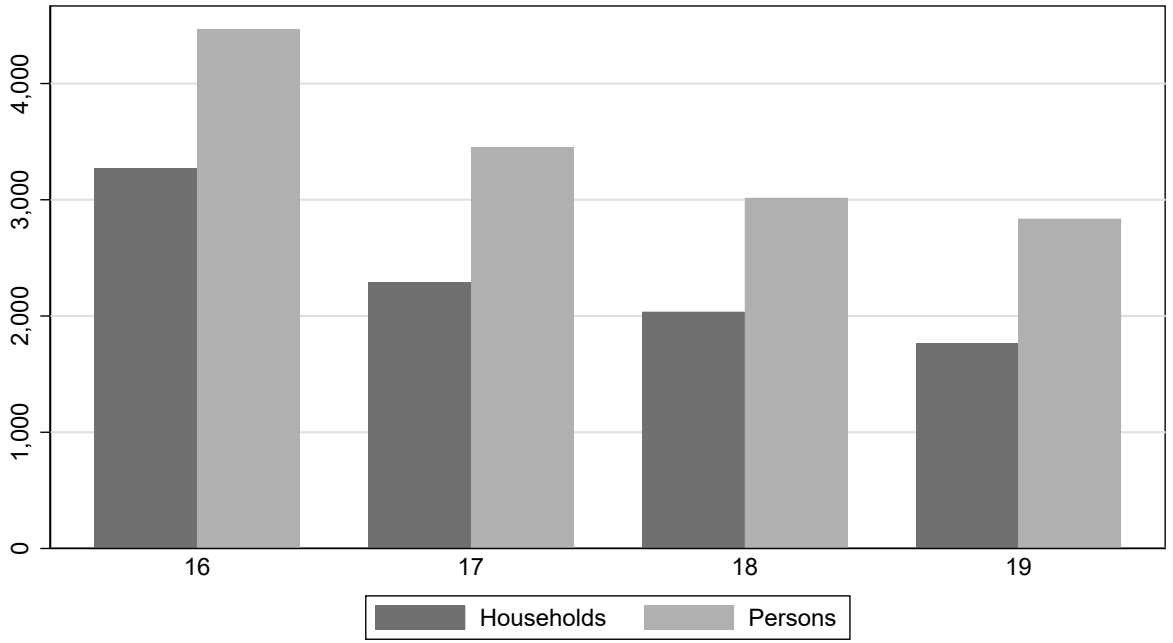

Figure 16: Comparison of Successful Interviews with Persons and Households (Subsamples M3/M4), Waves 1 to 4

Year	**2016**	**2017**	**2018**	**2019**
Persons	4,465	3,451	3,017	2,837
Households	3,273	2,291	2,037	1,764

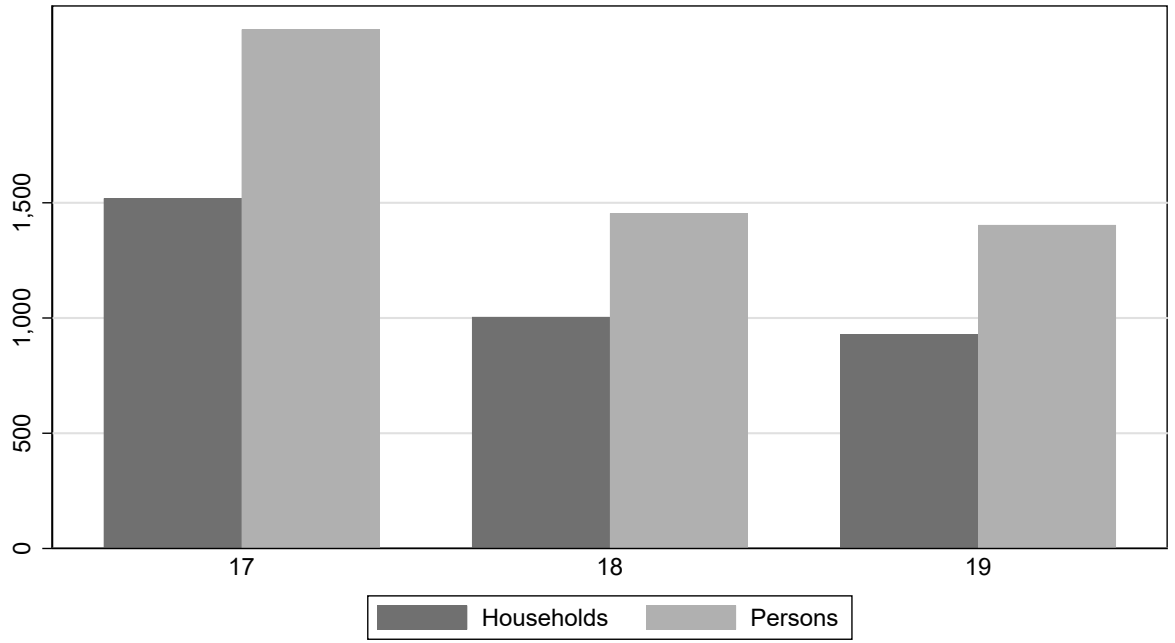

Figure 17: Comparison of Successful Interviews with Persons and Households (Subsample M5), Waves 1 to 3

Year	**2017**	**2018**	**2019**
Persons	2,252	1,454	1,404
Households	1,519	1,005	929

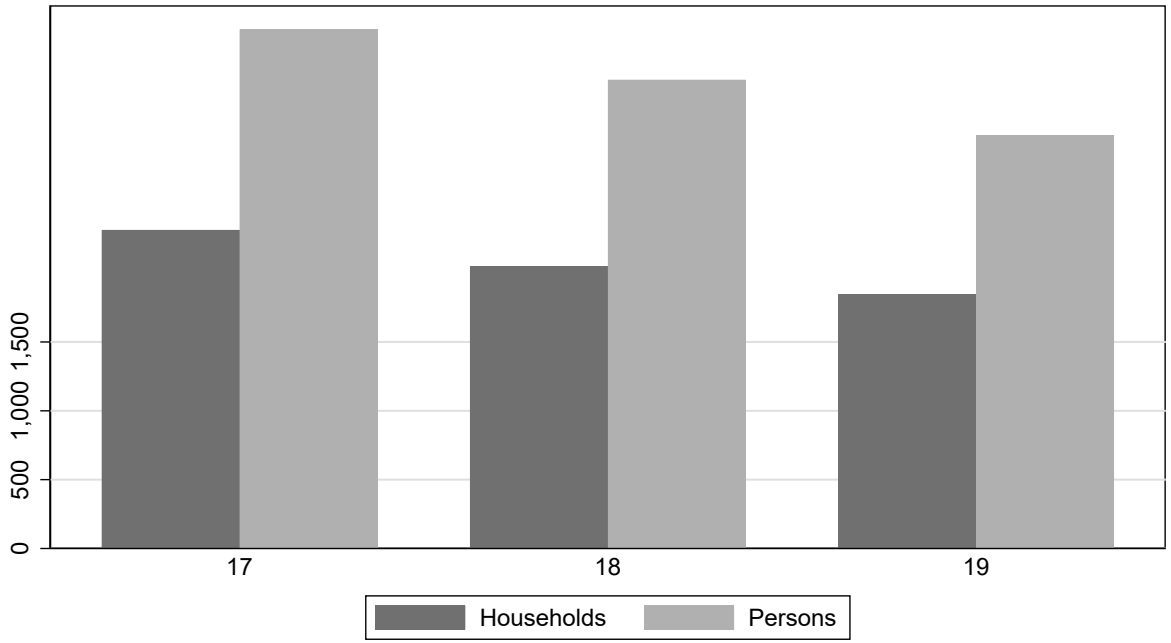

Figure 18: Comparison of Successful Interviews with Persons and Households (Subsample N), Waves 1 to 3

Year	**2017**	**2018**	**2019**
Persons	3,770	3,405	3,000
Households	2,314	2,114	1,889

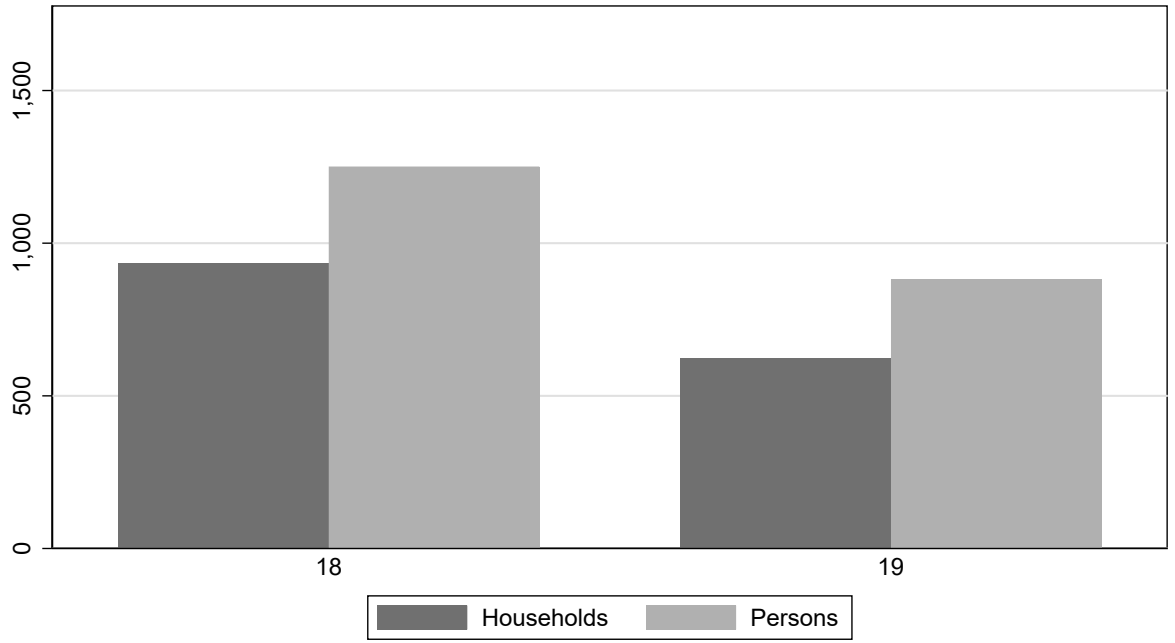

Figure 19: Comparison of Successful Interviews with Persons and Households (Subsample O), Waves 1 to 2

Year	**2018**	**2019**
Persons	1,251	882
Households	935	625

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

The following figures display the participation behavior of the first-wave respondents in the subsequent years distinguishing between continued participation (“With interview”), exits due to survey-unrelated attrition (“Moved abroad”, “Deceased”, “Under the age of 16”), and exits due to survey-related attrition (“Temporary drop-out”, “Drop-out”).

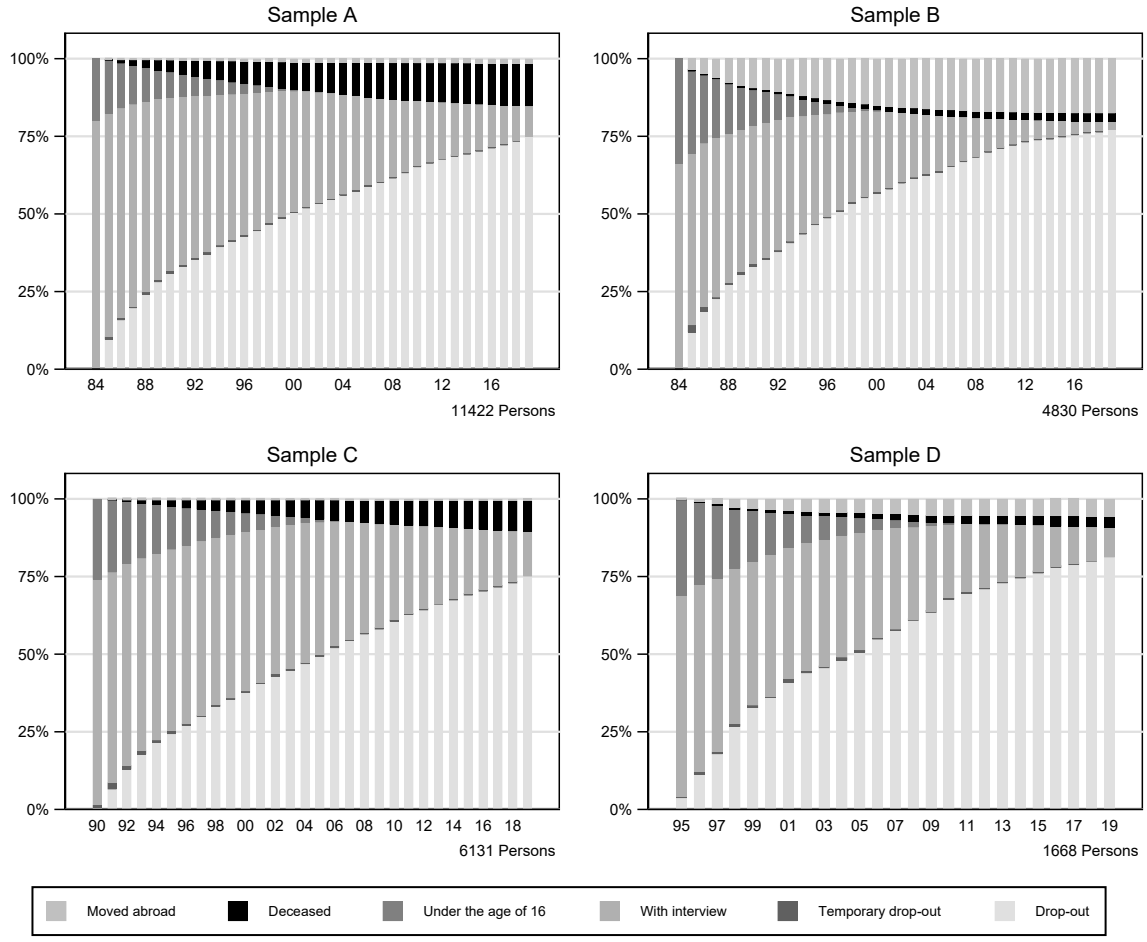

Figure 20: First-Wave Persons and their Participation Behavior. Development up to 2019

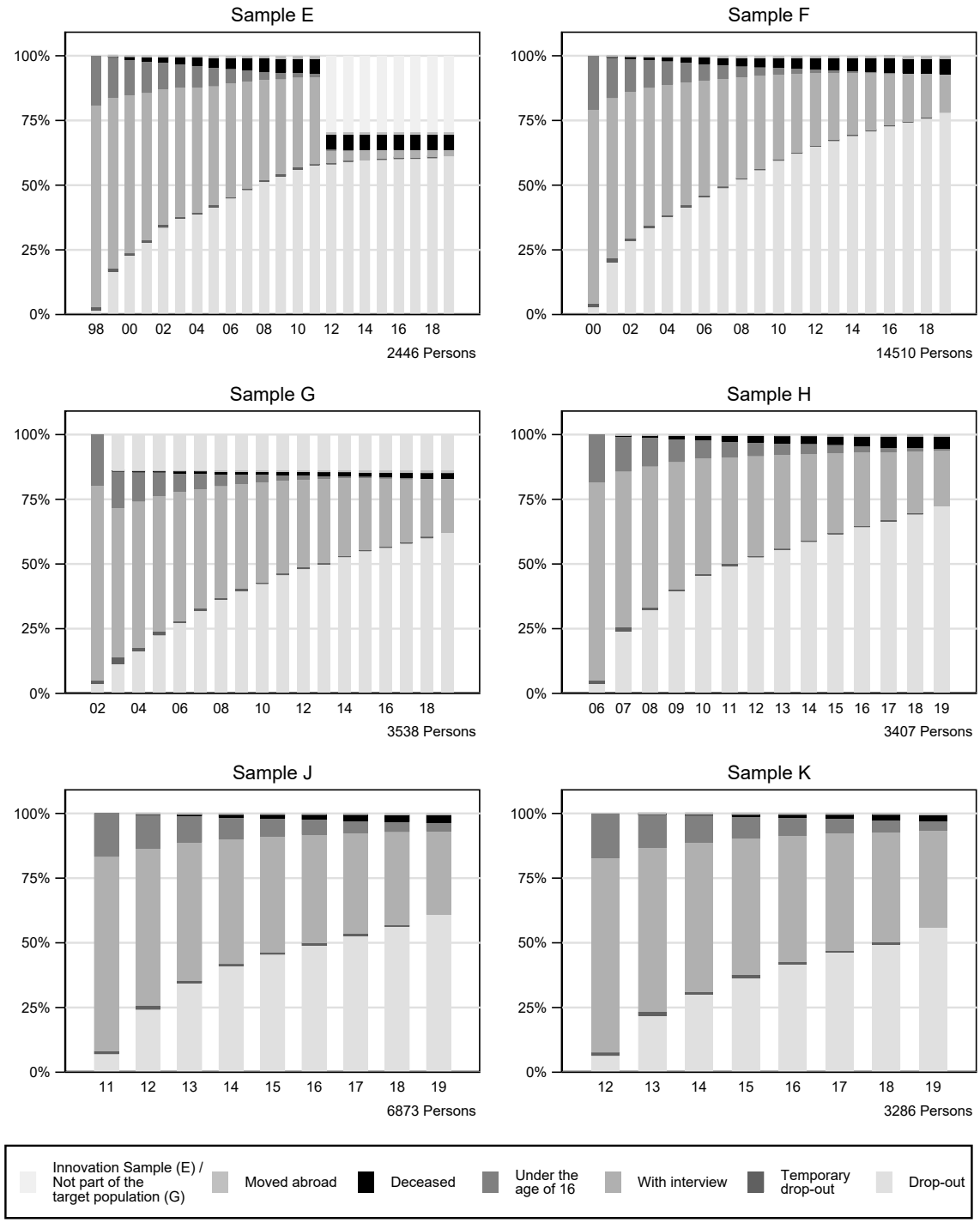

Figure 20: First-Wave Persons and their Participation Behavior. Development up to 2019

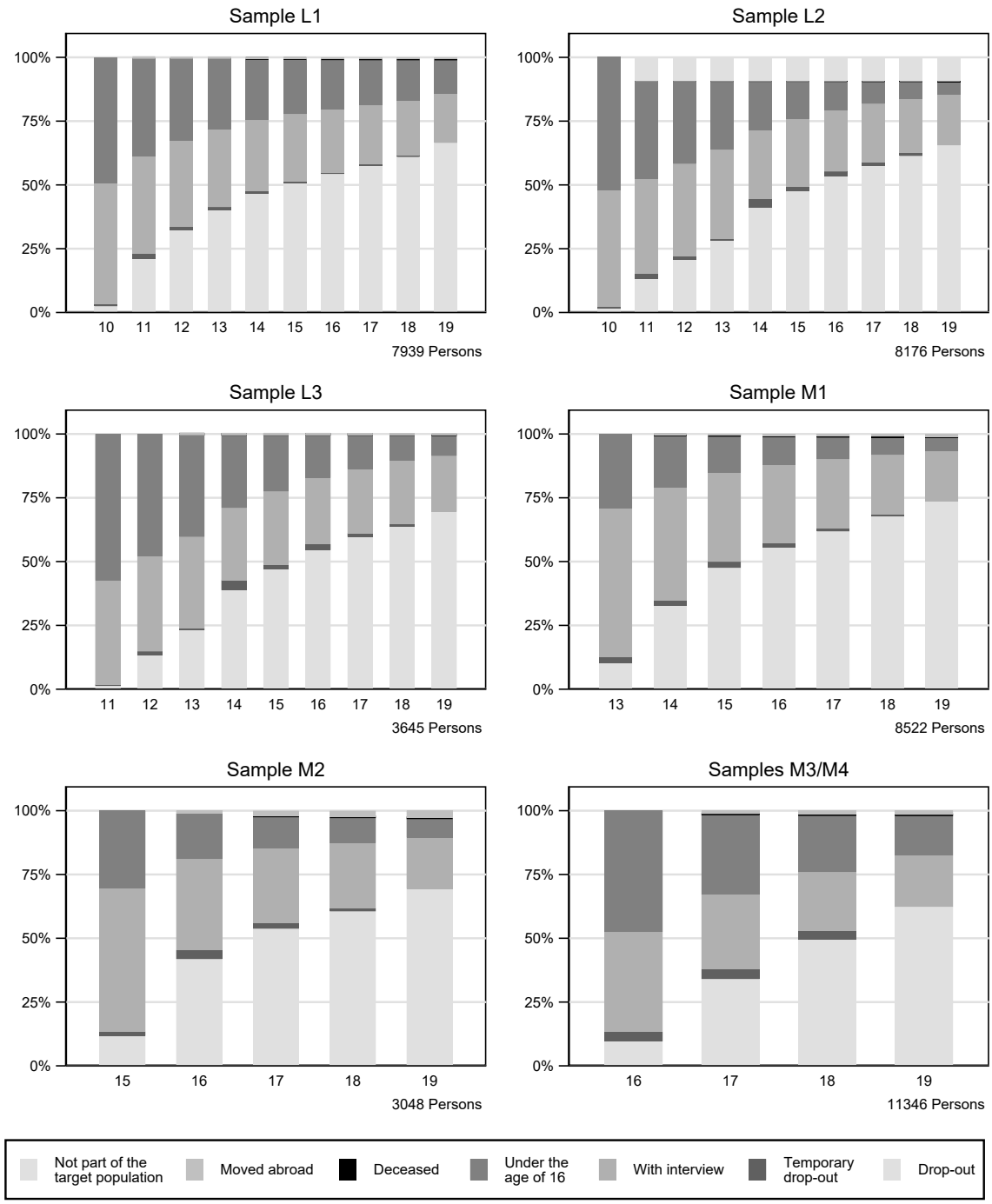

Figure 20: First-Wave Persons and their Participation Behavior. Development up to 2019

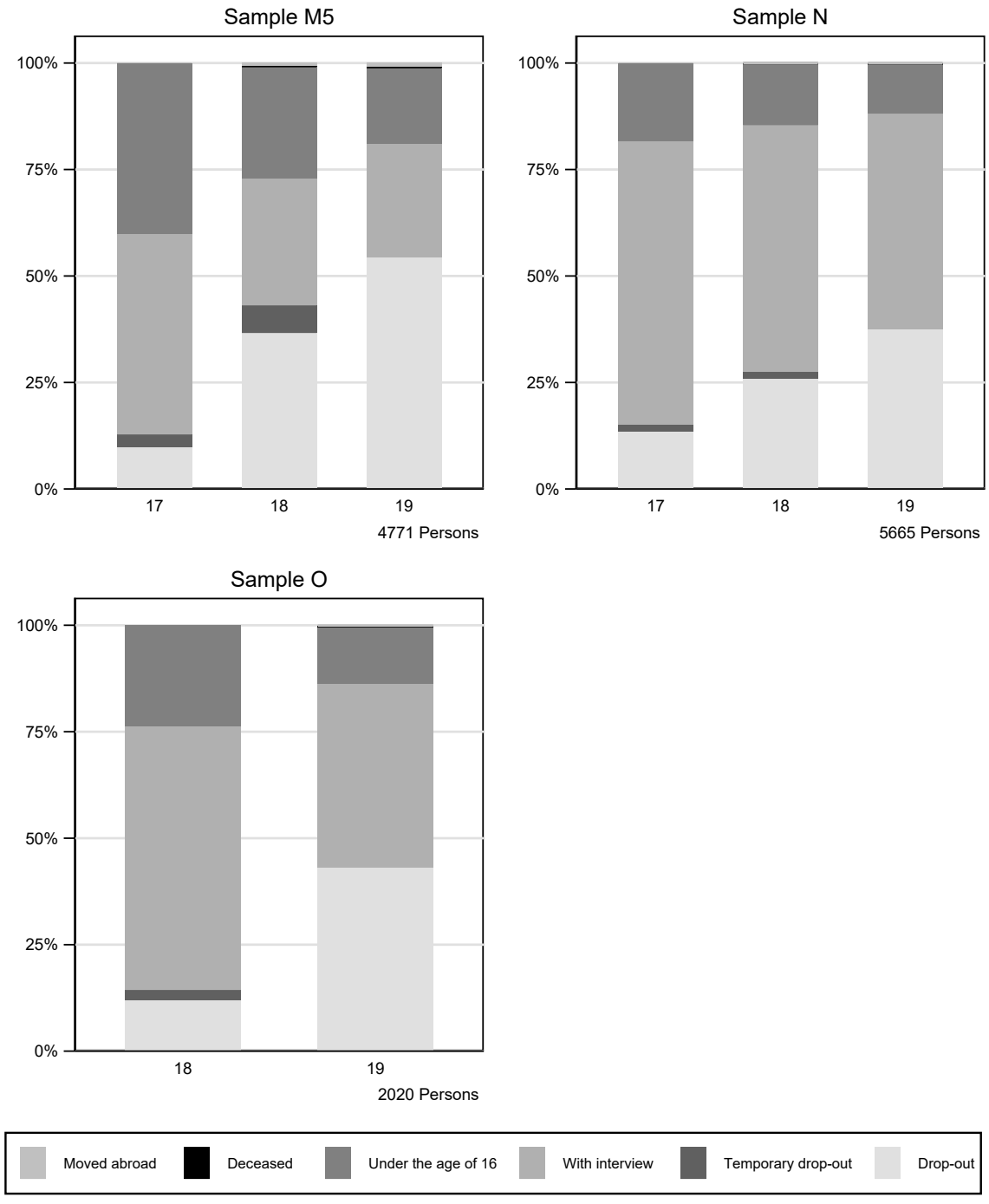

Figure 20: First-Wave Persons and their Participation Behavior. Development up to 2019

2.3 New Entrants through birth or move into SOEP Households and their Participation Behavior

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

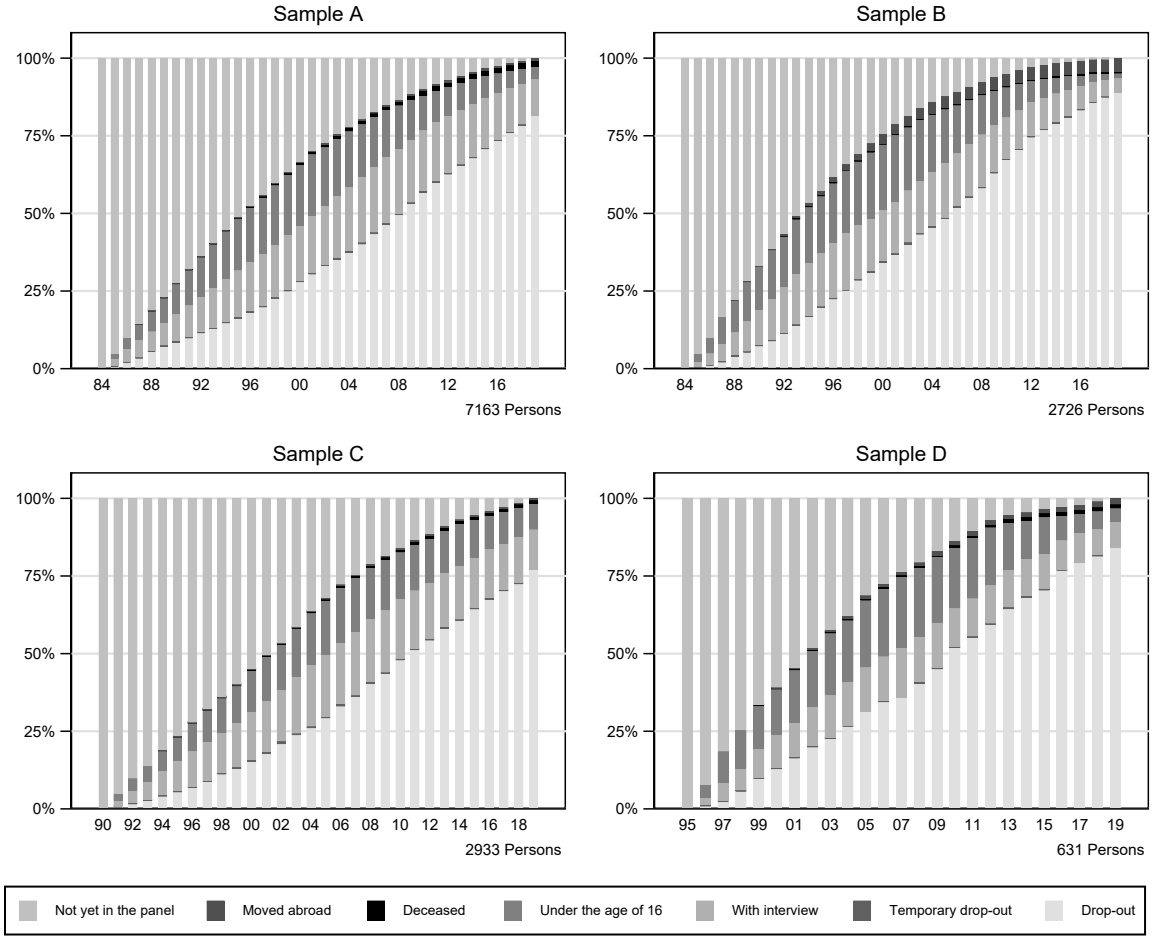

Figure 21: Entrants and their Participation Behavior. Development up to 2019

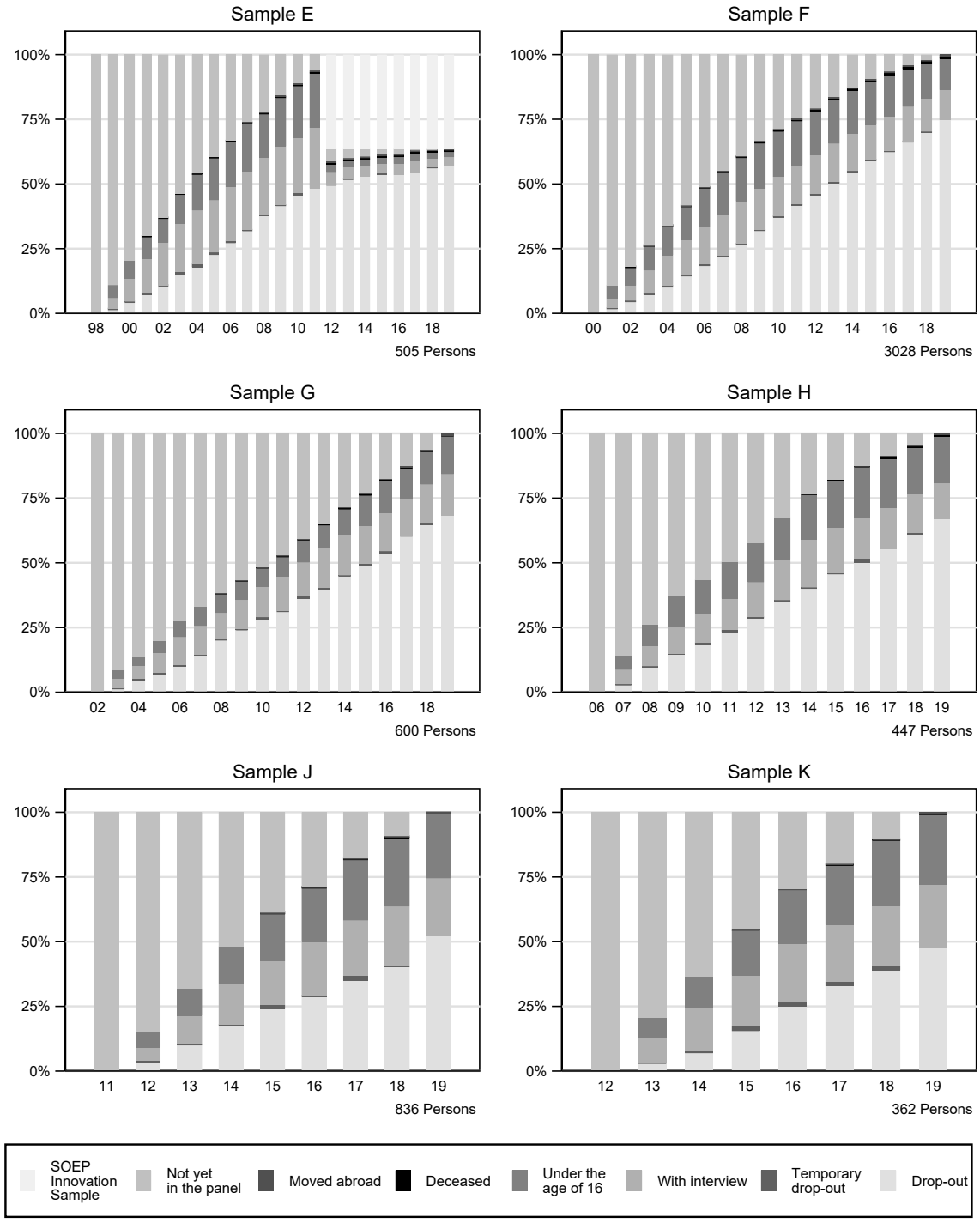

Figure 21: Entrants and their Participation Behavior. Development up to 2019

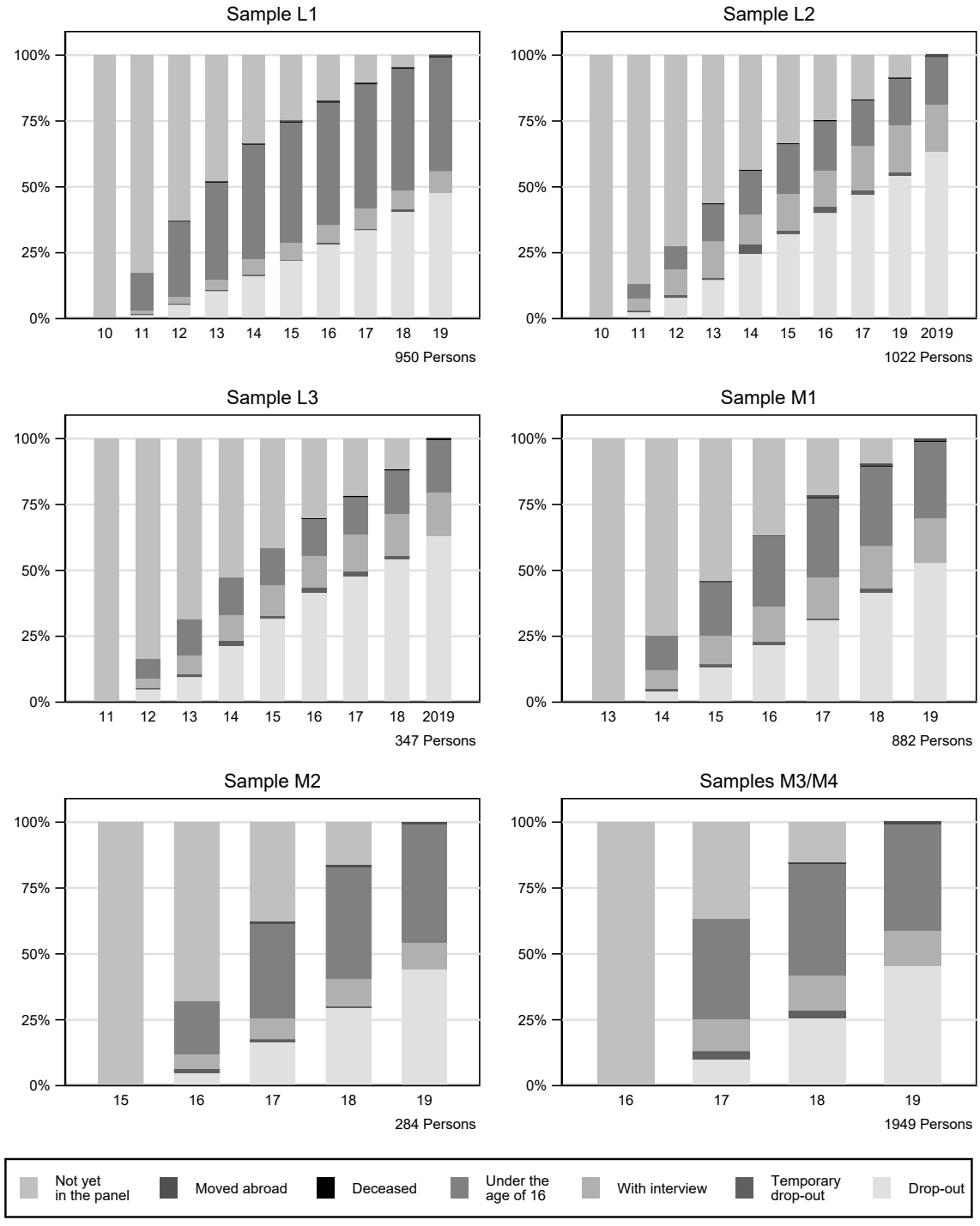

Figure 21: Entrants and their Participation Behavior. Development up to 2019

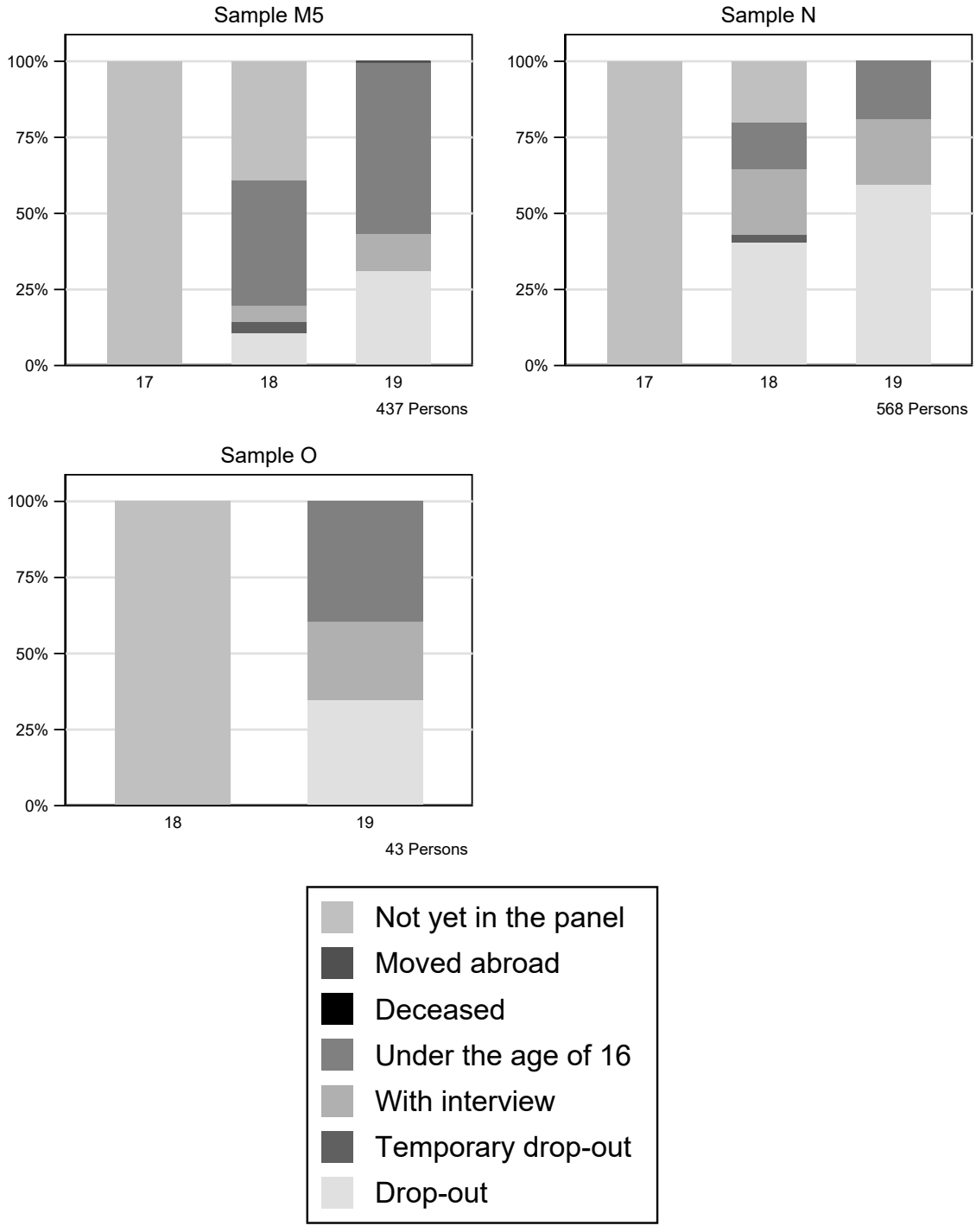

Figure 21: Entrants and their Participation Behavior. Development up to 2019

2.4 Original Households and Split-Offs

In case a household splits in multiple households (for instance, because a household member moves into another apartment), all resulting split-off households will be interviewed. The household which is not moving keeps the initial household number. These households are referred to as an “original household"[28]. The following figures display the development of the share of original households for each sample.

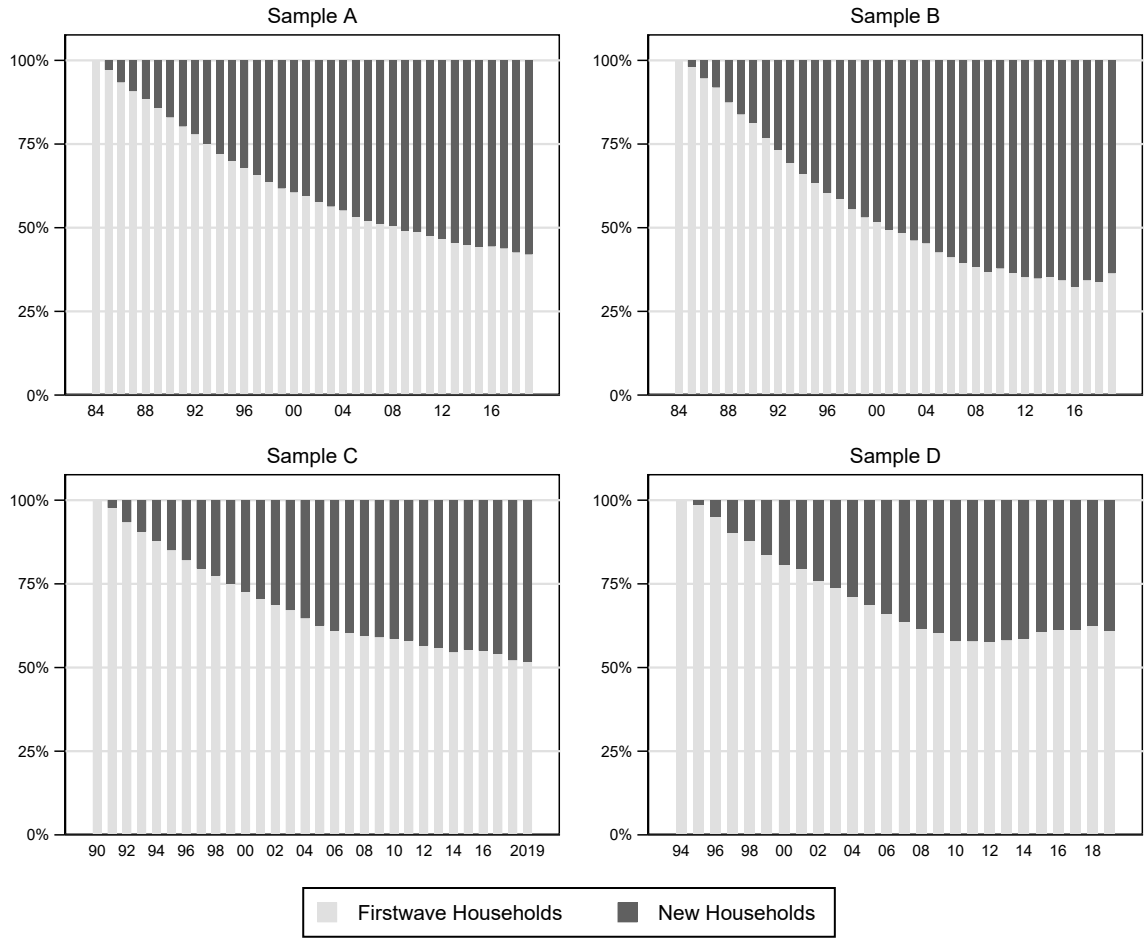

Figure 22: Proportion of First-Wave and New Households. Development up to 2019

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

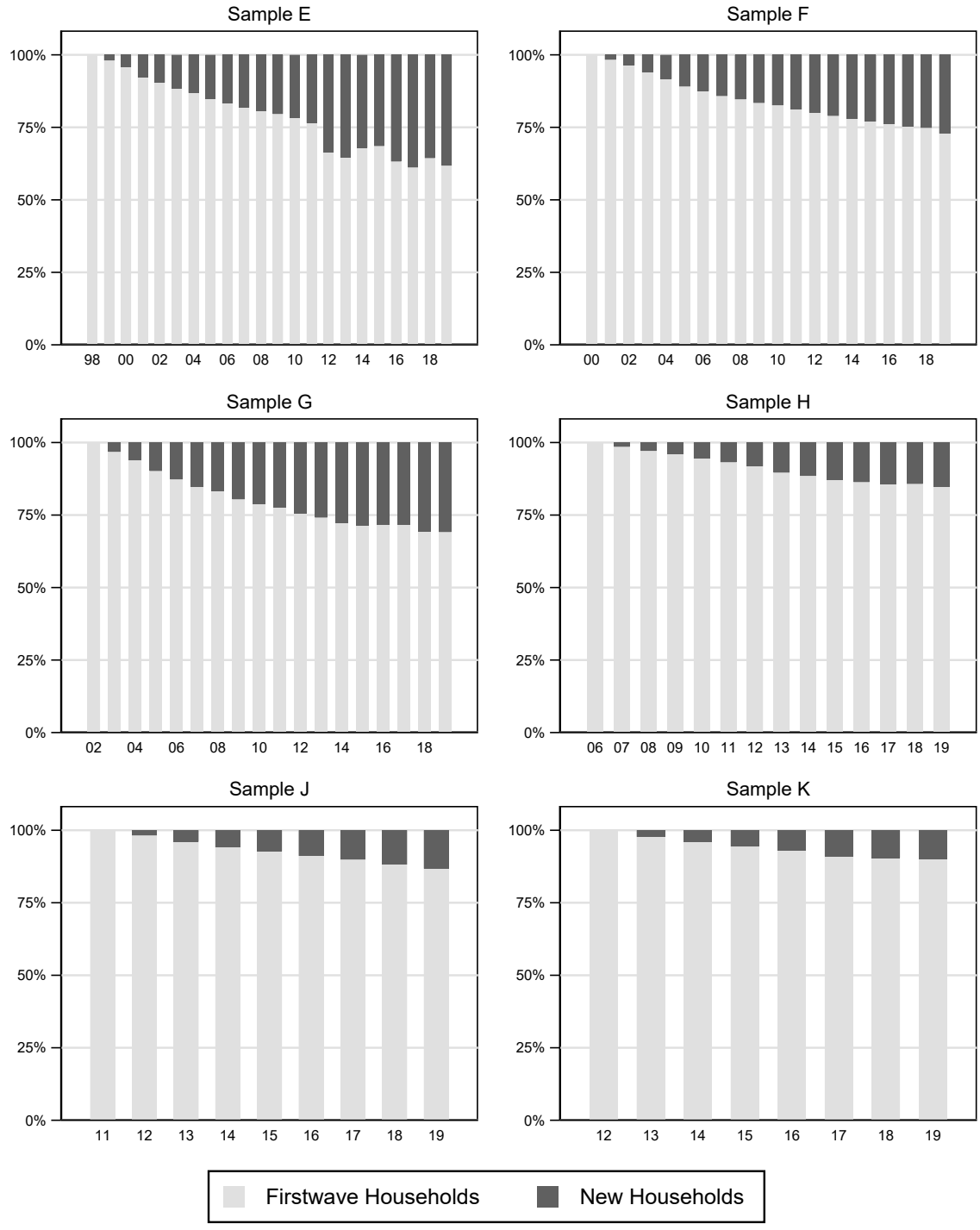

Figure 22: Proportion of First-Wave and New Households. Development up to 2019

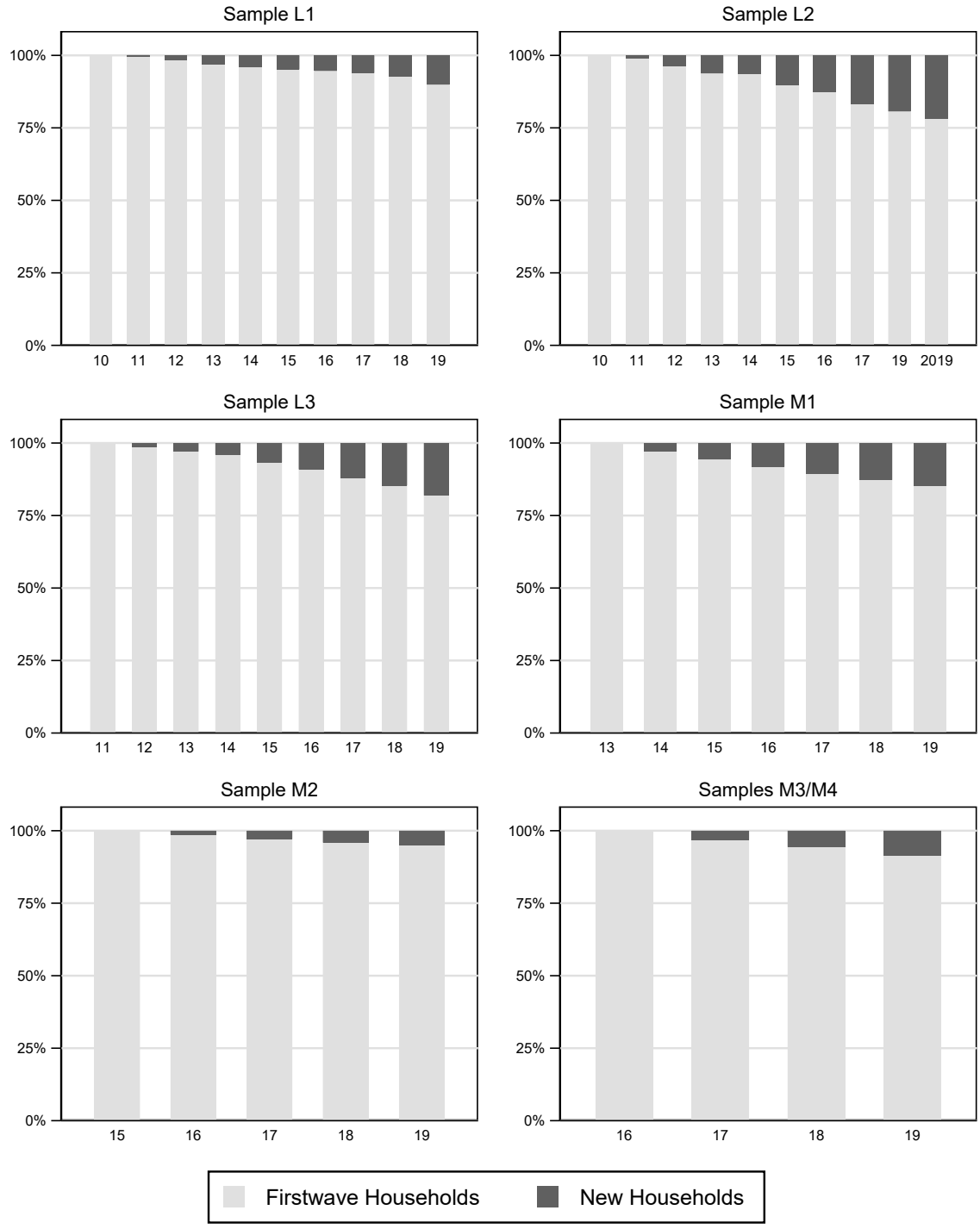

Figure 22: Proportion of First-Wave and New Households. Development up to 2019

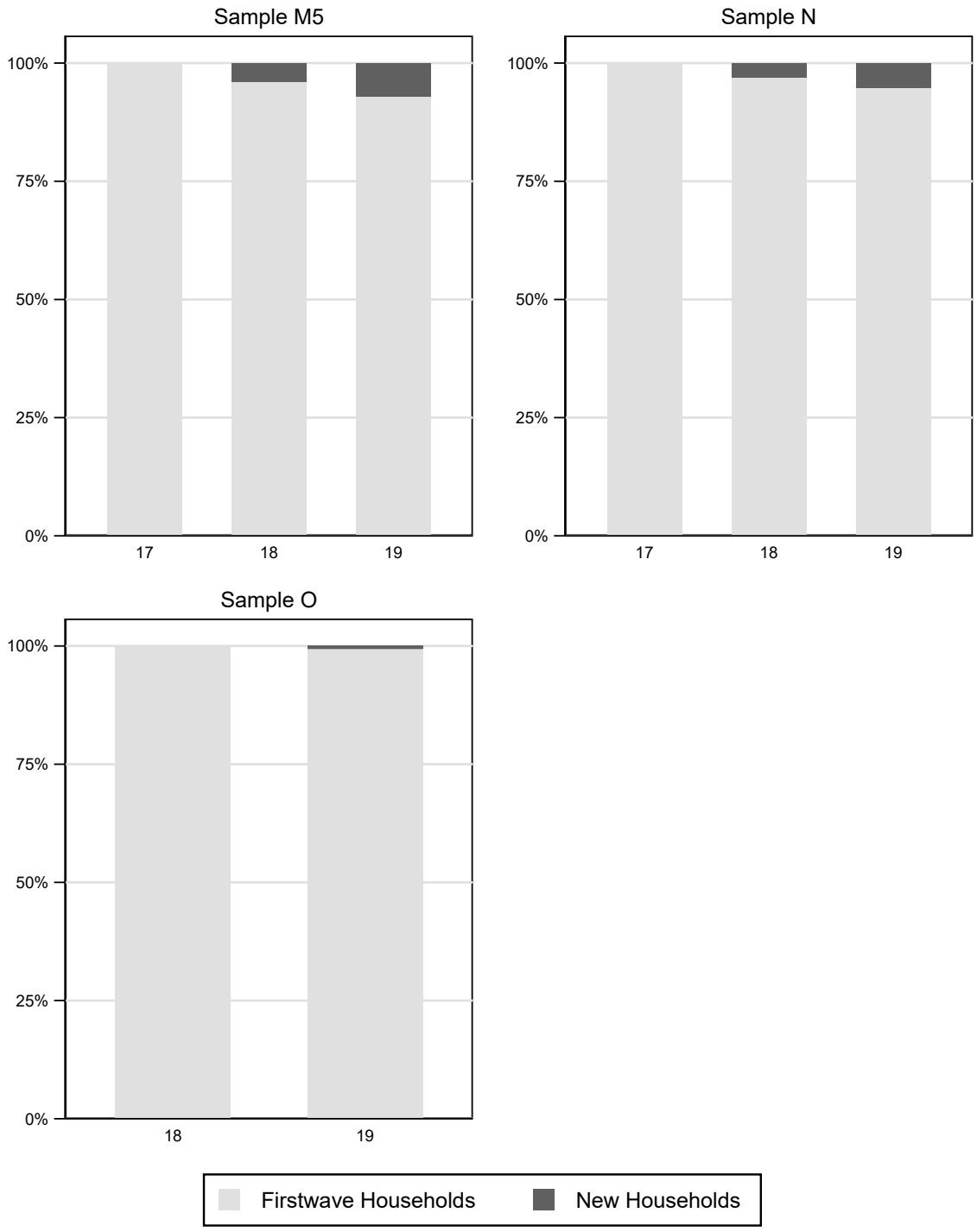

Figure 22: Proportion of First-Wave and New Households. Development up to 2019

2.5 The Risk of Survey-Related Panel Attrition

The following figures display Kaplan-Meier estimates of the survey related attrition risk (unsuccessful follow-up and refusal) of the net sample of first-wave respondents thereby ignoring survey unrelated exits (moves abroad and deaths). These figures stratify the drop-out risk in different groups of the sample defined by respondents' sample membership (Figures 23 through 27) and some basic socio-demographic characteristics measured in the year of sampling, such as age, occupation, income, and education (Figures 29, 30, 31, and 32 respectively). 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 23), 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.

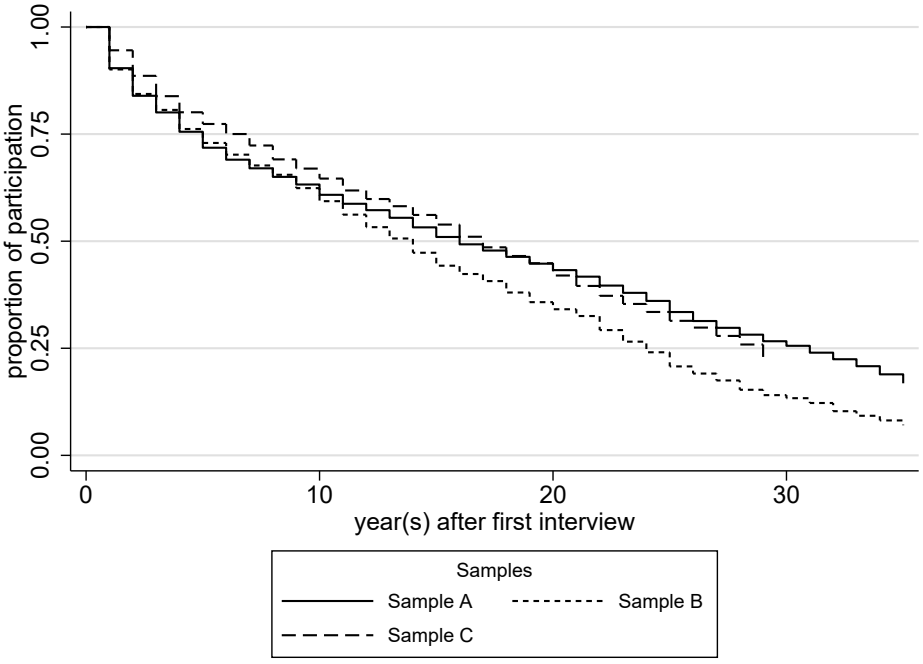

Figure 23: Successful Re-Interviewing of First-Wave Respondents by Subsamples A, B, C. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

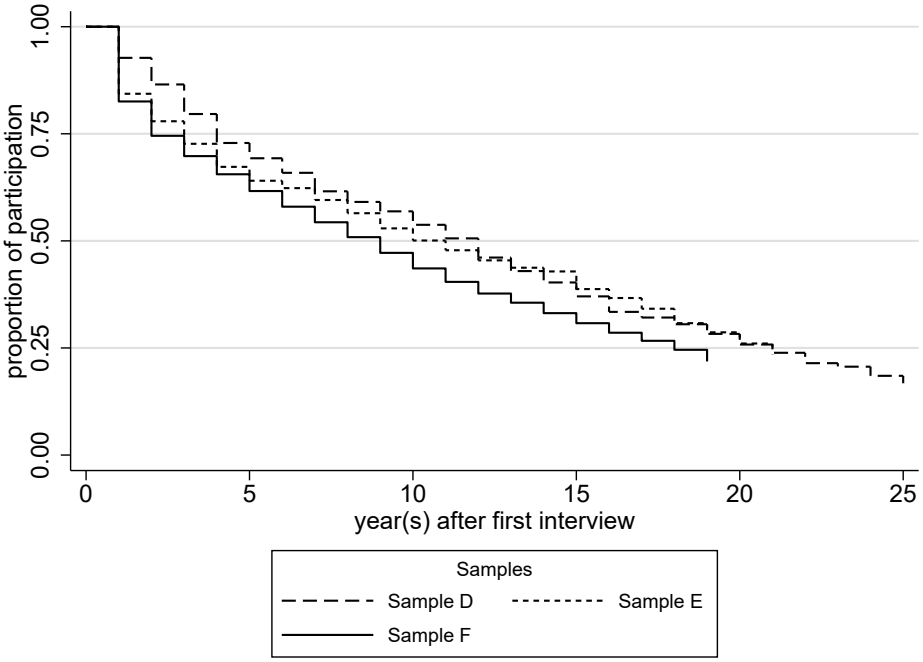

Figure 24: Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

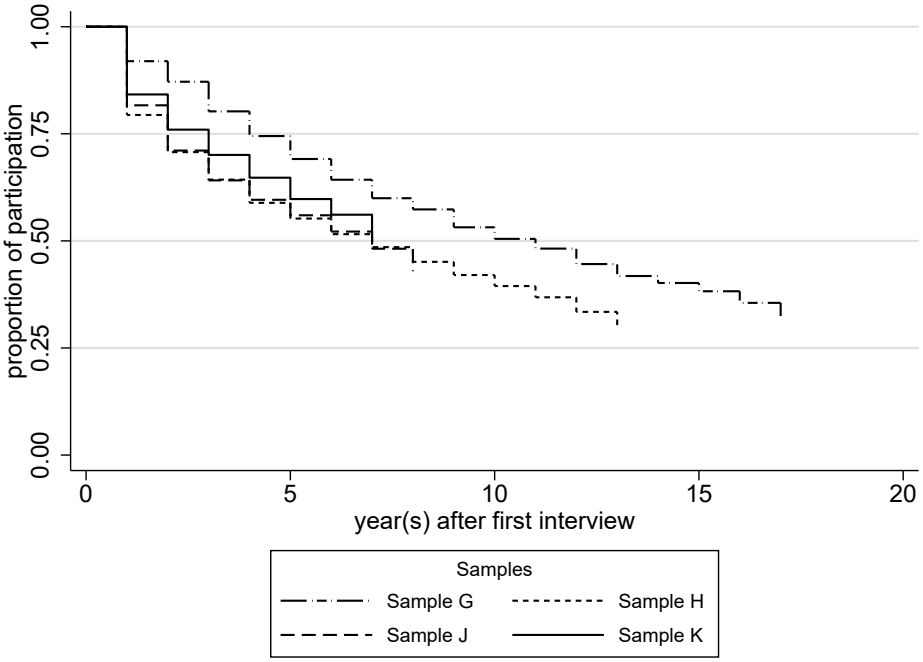

Figure 25: Successful Re-Interviewing of First-Wave Respondents by Subsamples G, H, J and K. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

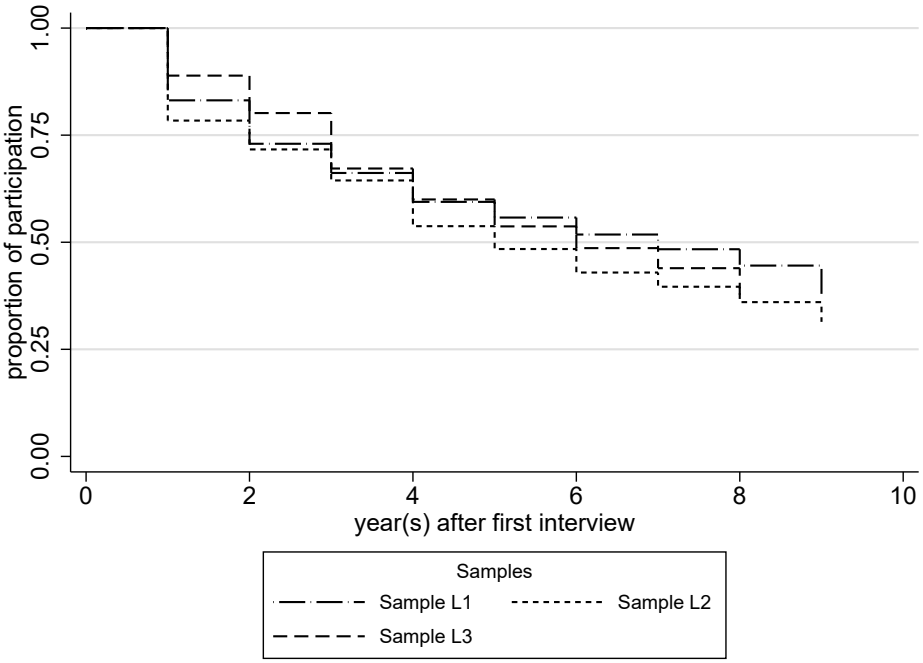

Figure 26: Successful Re-Interviewing of First-Wave Respondents by Subsamples L1, L2 and L3. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

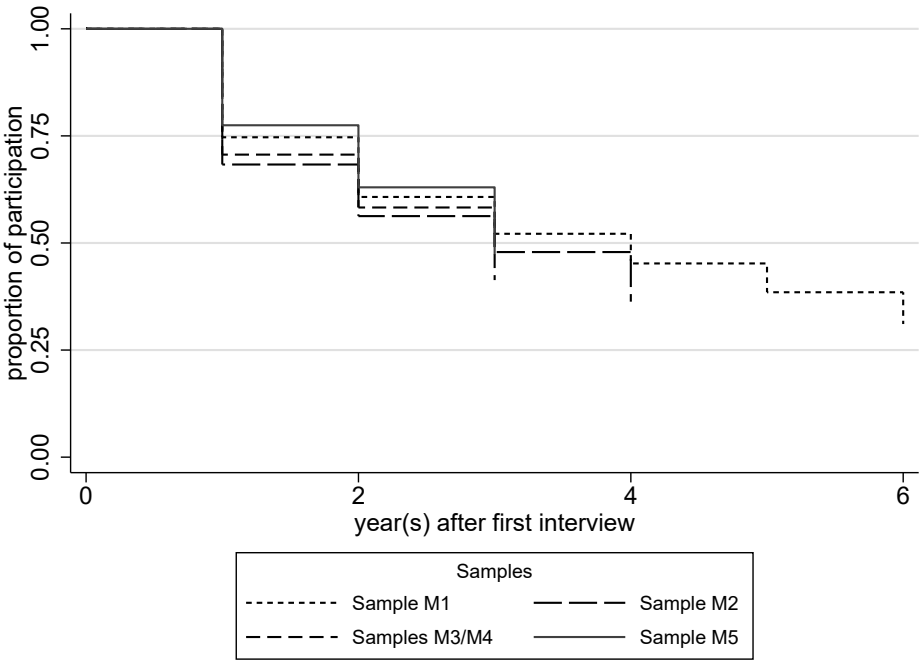

Figure 27: Successful Re-Interviewing of First-Wave Respondents by Subsamples M1, M2, M3/M4 and M5. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

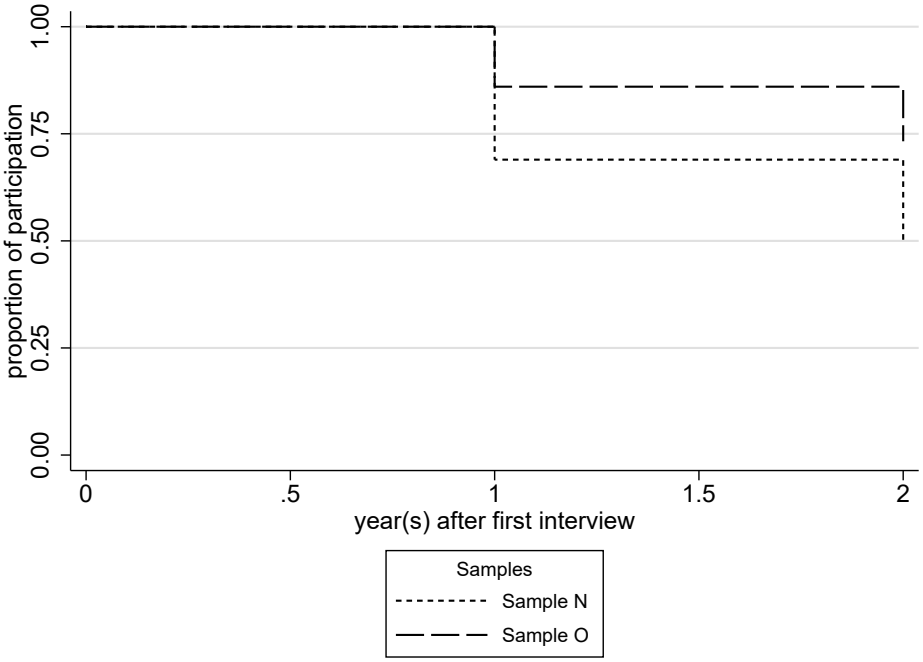

Figure 28: Successful Re-Interviewing of First-Wave Respondents by Subsamples N and O. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

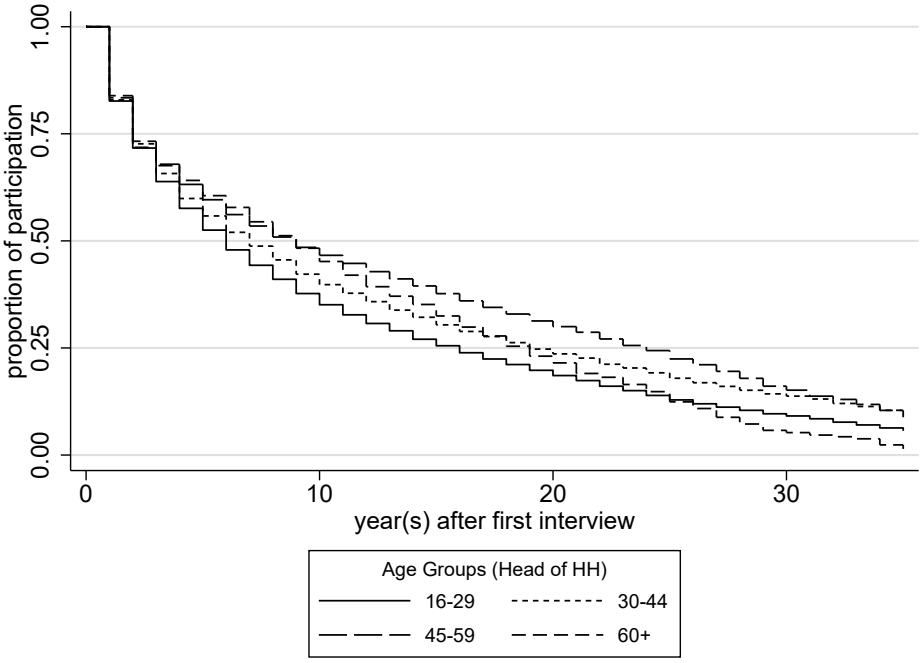

Figure 29: Successful Re-Interviewing of All First-Wave Respondents by Age Categories. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

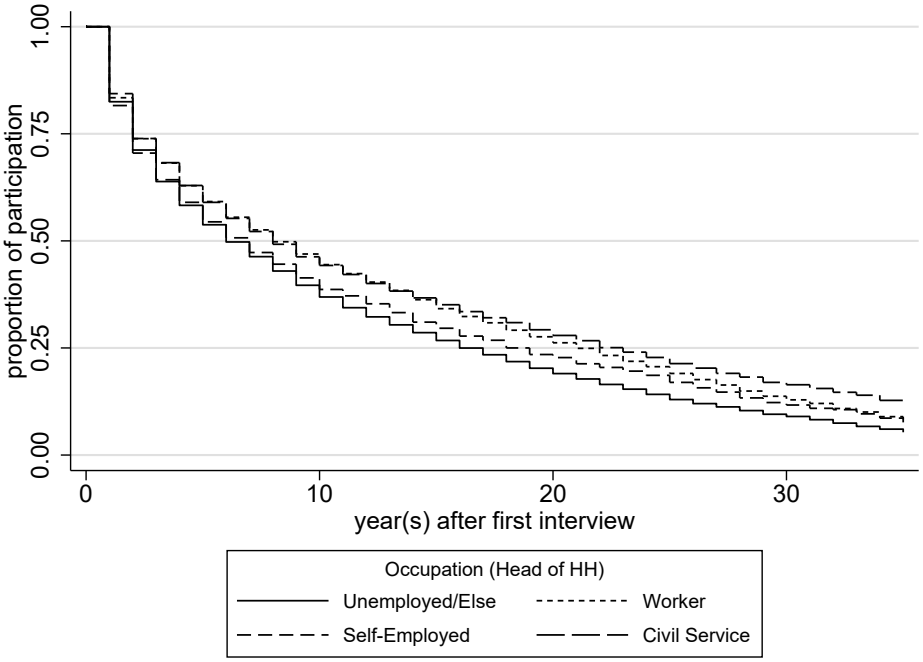

Figure 30: Successful Re-Interviewing of All First-Wave Respondents by Occupation. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

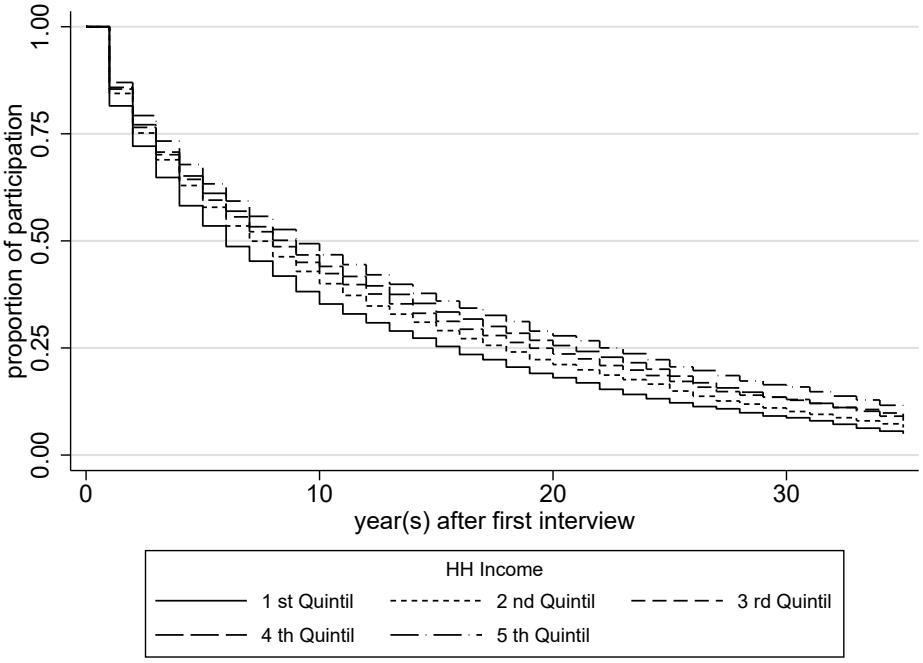

Figure 31: Successful Re-Interviewing of All First-Wave Respondents by Income Quintiles. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

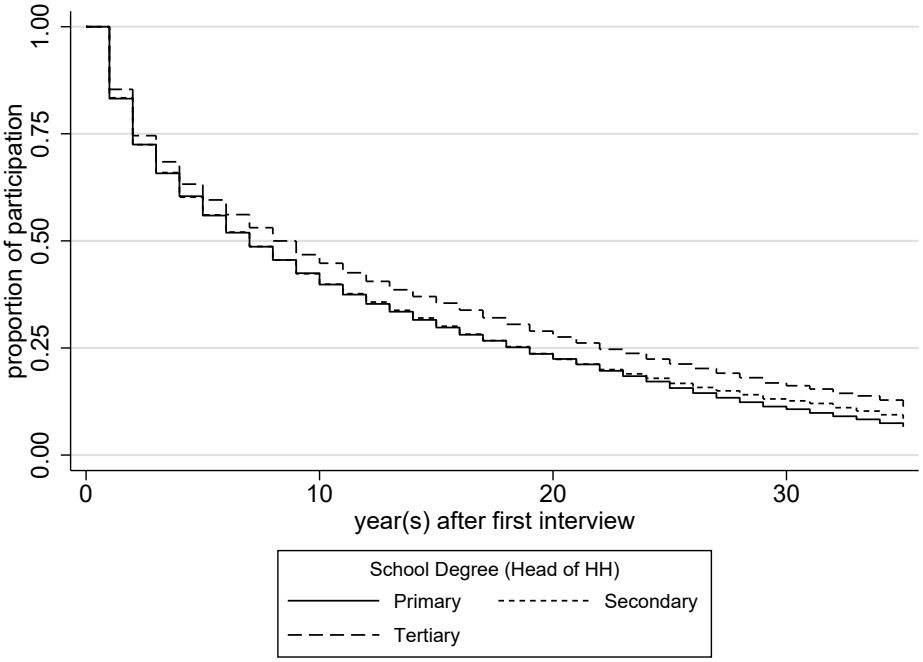

Figure 32: Successful Re-Interviewing of All First-Wave Respondents by Education. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad

3 Panel Attrition Due to Unsuccessful Follow-Ups

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

3.1 The Frequency of Successful Follow-Ups

Table 3.1 displays the number of households of the previous waves that need to be re-contacted and the relative frequency of successful follow-ups in subsamples A through Q and waves 1985 through 2019. 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.

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

Table 3.1: The Frequency of Households to be Re-Contacted and the Percentage of Successful Follow-Ups, Subsamples A to Q 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	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1984	4,528		1,393																			
1985	4,681	98.5	1,370	96.8																		
1986	4,486	99.0	1,325	97.4																		
1987	4,232	99.1	1,220	98.7																		
1988	4,140	99.2	1,191	99.1																		
1989	3,984	99.1	1,157	99.0																		
1990	3,902	99.2	1,124	98.8	2,179																	
1991	3,860	99.5	1,151	99.3	2,246	98.5																
1992	3,845	99.7	1,153	99.2	2,302	99.5																
1993	3,867	99.3	1,172	98.6	2,227	99.1																
1994	3,849	99.3	1,150	99.0	2,134	99.4	236															
1995	3,784	99.5	1,108	99.0	2,110	99.6	540	100.0														
1996	3,747	99.7	1,069	99.3	2,103	99.5	544	99.6														
1997	3,688	99.6	1,038	99.1	2,087	99.5	541	99.3														
1998	3,667	99.4	1,019	99.4	2,079	99.4	528	99.1	1,056													
1999	3,631	99.6	975	99.4	2,037	99.7	498	99.4	1,089	99.5												
2000	3,549	99.6	934	99.5	2,025	99.7	467	99.8	967	99.2	6,043											
2001	3,463	99.6	904	99.4	2,034	99.7	454	99.1	921	99.1	6,162	99.0										
2002	3,406	99.7	877	99.1	2,005	99.6	450	99.8	873	99.4	5,447	99.5	1,224									
2003	3,330	99.6	840	99.6	1,982	99.6	434	99.5	834	99.3	4,965	99.7	1,056	99.1								
2004	3,260	99.8	803	99.6	1,962	99.6	436	99.8	797	99.7	4,736	99.6	1,010	99.7								
2005	3,220	99.8	779	99.4	1,959	99.7	429	99.3	783	99.9	4,577	99.7	1,001	99.7								
2006	3,138	99.7	770	99.6	1,941	99.4	425	98.8	775	99.1	4,401	99.3	995	99.5	1,506							
2007	3,000	99.7	725	99.4	1,834	99.9	387	99.5	727	99.7	4,157	99.5	933	99.2	1,530	99.5						
2008	2,856	99.8	676	99.3	1,767	99.5	372	99.5	680	99.7	3,962	99.4	904	99.7	1,326	99.6						
2009	2,730	99.7	620	99.4	1,695	99.9	351	99.7	636	100.0	3,760	99.6	870	99.5	1,145	99.7	1,495					
2010	2,570	99.8	548	99.5	1,627	100.0	334	99.7	605	99.8	3,538	99.6	826	99.9	1,059	99.5	1,738	98.3				
2011	2,421	99.8	495	99.2	1,541	99.8	303	99.3	589	100.0	3,319	99.7	797	99.6	992	99.6			3,136			
2012	2,289	99.8	440	99.8	1,466	99.9	286	100.0	116	99.1	3,076	99.9	774	99.7	928	99.9			3,204	99.2	1,526	
2013	2,180	99.6	393	99.2	1,417	99.7	269	99.3	98	100.0	2,881	99.7	733	99.6	877	99.5			2,871	99.5	1,564	99.0
2014	2,078	99.4	361	99.4	1,351	99.6	249	100.0	90	100.0	2,741	99.7	725	99.3	828	99.4			2,519	99.1	1,448	99.4
2015	1,998	99.4	331	99.4	1,300	99.5	229	100.0	83	100.0	2,597	99.2	699	99.3	790	99.7			2,309	99.4	1,308	99.3
2016	1,861	99.6	296	99.7	1,217	99.7	208	99.5	83	96.4	2,412	99.4	669	98.7	720	99.6			2,119	99.5	1,209	99.3
2017	1,748	99.3	271	98.9	1,125	99.6	184	99.5	75	100.0	2,214	99.3	622	99.5	677	99.4			2,015	99.3	1,105	99.5
2018	1,641	99.4	236	100.0	1,060	99.2	174	99.4	69	100.0	2,070	99.4	608	99.0	641	99.7			1,916	99.3	1,067	99.0
2019	1,497	99.6	200	99.5	988	99.8	155	100.0	64	100.0	1,947	99.6	574	99.8	593	99.3			1,797	99.2	987	99.4

Year	Sample L1		Sample L2		Sample L3		Sample M1		Sample M2		Sample M3/4		Sample M5		Sample N		Sample O		Sample P		Sample Q	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
2010	2,074		2,500																			
2011	2,083	98.8	2,271	98.1	924																	
2012	1,867	99.6	2,255	98.5	943	98.7																
2013	1,753	99.3	2,177	98.8	920	99.1	2,723															
2014	1,512	99.4	2,027	98.2	836	98.6	2,828	98.8														
2015	1,404	99.4	1,880	98.4	789	97.8	2,456	98.0	1,096													
2016	1,287	99.5	1,736	98.6	732	98.2	2,116	97.7	1,096	97.1	3,289											
2017	1,209	99.3	1,587	98.6	686	97.8	1,794	97.9	931	98.1	3,351[29]	93.6	1,519		2,314							
2018	1,151	98.9	1,494	98.1	650	98.2	1,605	98.4	688	98.7	3,110	94.0	1,585	93.2	2,482	99.1	935					
2019	1,086	99.4	1,451	97.7	621	97.3	1,415	98.2	580	96.7	2,618	93.2	1,491	88.3	2,335	99.3	944	98.1	1,960		477	

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

3.2 Predicting the Probability of Successful vs. Unsuccessful Follow-Ups in the Year 2019

Based on household and interview level characteristics measured in 2018, we aim to predict the probability of re-contacting a household relative to unsuccessful follow-up in 2019. Among a very large number of regressors that we tested in preliminary analyses, we identified a small set of variables that exert a robust effect on the probability of successful follow-ups ($p < 0.05$). Table 3.2 describes the regressors and Table 3.3 reports the subsample-specific estimates of logit models for the probability of re-contacting a household relative to unsuccessful follow-up.

Note that the estimates of regression models of the previous waves from 1985 to 2018 are not reported in the present data documentation due to space restrictions. These can be obtained from previous attrition documentations (e.g. Siegers et al. (2020)).

Table 3.2: Definition of the Regressors of the Logit Model of Refusal

Variable	Label	Value
Interview Characteristics		
New Household	Household new in SOEP	0/1
New Address	Household moved	0/1
New SOEP Member	Head of household has had less than 4 interviews	0/1
SOEP Household Moved	Participating household moved to new location	0/1
Temporary Drop-Out	Temporary drop-out of household in previous year	0/1
Temp. Drop-Out Related HH	Temporary drop-out of related household	0/1
Email Known	Email address disclosed	0/1
Phone Unknown	Telephone number undisclosed	0/1
Change in HH Composition	Recent change in household composition	0/1
Change of Interviewer	Latest interview conducted by different interviewer	0/1
Person Present Not Living in HH	Person Present during interview who does Not Live in household	0/1
(High) Item Nonresponse HH	(High) Item Nonresponse household	0/1
Demographic Characteristics		
Family Household	Family household	0/1
Single Household	One-person household	0/1
Female Head of HH	Head of household is female	0/1
Younger than 25	Head of household is younger than 25 years old	0/1
Separation	Household member(s) separated from partner last year	0/1
Work, Education, and Finances		
Same Employer 3rd Q.	Head of household with current employer since third quarter	0/1
Same Employer 4th Q.	Head of household with current employer since fourth quarter	0/1
Left Job Prev. Year	Head of household left job at the beginning of the previous year	0/1
In Education	Head of household is currently pursuing education	0/1
Vocational Training	Head of household completed vocational training previous year	0/1
Blue-Collar Worker	Head of household is a blue-collar worker	0/1
Unemployment benefits	Head of household currently receives unemployment benefits (ALG II)	0/1
Health, Personality, and Activities		
Individual Health Services	Head of household received health care not covered by insurance	0/1
Low Life Satisfaction	Head of household is dissatisfied with his/her life	0/1
Very energetic	Head of household often feels very energetic	0/1
Achieved Less	Head of household achieved less than desired due to emotional problems	0/1

Table 3.2 – *Continued from previous page*

Variable	Label	Value
Vegetarian/vegan	Head of household does not consume meat	0/1
Middle-class	MICROM Sinus-Geo-Milieu im Haushalt: "Bürgerliche Mitte"	0/1
Sustainability and Social Justice	MICROM Sinus-Geo-Milieu im Haushalt: "Sozialökologisch ausgerichtet"	0/1
Outsider Feeling	Head of household often feels like an outsider	0/1
Political Party Preference	Head of household has a general party preference	0/1
Many Friends	Head of household has many close friends	0/1
No Concerns About Xenophobia	Head of household has no concerns about xenophobia in the general population	0/1
Worried About Own Economic Situation	Head of household often worries about own financial/economic situation	0/1
Building, Area, and Region		
Many Single-Households	Area with many single households	0/1
No balcony/terrace	No balcony or terrace in dwelling	0/1
No garage	No garage in dwelling	0/1
No Car in HH	No household member owns a car	0/1
High-Rise Area	Household located in area with large number of high-rises	0/1
New Apartments	High share of newly constructed apartments in one- and two-family houses	0/1
Manufacturing Sector Low	Low share of persons employed in the manufacturing sector	0/1
Minijobs Low	Low share of minijobs among all types of employment	0/1
Fathers Parental Allowance	High share of fathers receiving parental allowance	0/1
Apprentices Low	Low number of apprentices per 100 residents 15-25 years old	0/1
Employment Rate High	Household located in area with high employment rate	0/1
Fewer Women Employed	Proportion of employment rate of women to men – 1st quartile	0/1
Commute 300km and more 4th quartile	Household located in area with high number of commutes of 300km and more	0/1
Internal Migration Low	Household located in area with low internal migration net total	0/1
Development Employment High	Area's employment development – 4th quartile	0/1
Development Tax Revenue High	City's tax revenue development – 4th quartile	0/1
Migration net total 1st quartile	Household located in area with low net total migration	0/1
Life Expectancy Low	Household located in area with low life expectancy	0/1
Schleswig-Holstein/Hamburg	Household located in Schleswig-Holstein/Hamburg	0/1
City 20,000 - 100,000 residents	Household located in city with 20,000 - 100,000 residents	0/1
High Share of AfD Voters	Household located in area with high share of AfD voters	0/1
Low Share of CDU/CSU Voters	Household located in area with low share of CDU/CSU voters	0/1
High Share of Green Party Voters	Household located in area with high share of Green Party voters	0/1
High Share of Wind Energy	Household in area with high share of wind energy	0/1

Table 3.3: Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2019

Explanatory Variable	**Sample A**	**Sample F**	**Sample H**	**Sample J**	**Sample K**	**Sample L1**	**Sample L2**	**Sample L3**	**Sample M1**	**Sample M2**	**Sample M3/4**	**Sample M5**	**Sample N**	**Sample O**
Intercept	2.87***	2.00***	2.18***	2.87***	2.21***	2.00***	3.57***	2.32***	2.97***	4.65***	1.31***	2.62***	4.18***	4.31***
Interview Characteristics														
New Household		−0.91***	−1.42**	−0.69*	−1.34***	−0.73*	−0.89**	−0.67*	−1.33***	−3.28**	−0.89***	−1.36***	−1.61***	
New Address						−0.86***	−1.56***		−1.47***	−3.37***	−1.54***	−1.65***	−1.87***	−2.56***
New SOEP Member		−1.03**												
SOEP Household Moved								−0.88*						
Temporary Drop-Out							−1.14***					−0.80***		
Temp. Drop-Out Related HH				−0.99**										
Email Known												0.58*		
Phone Unknown				−0.83**			−1.49***		−0.83**	−1.54**	−0.56***	−0.96***	−1.08***	−0.85*
Change in HH Composition							−0.77**							
Change of Interviewer	−1.17*													
Person Present Not Living in HH														−1.41**
(High) Item Nonresponse HH									−0.55*					
Demographic Characteristics														
Family Household											0.25*			
Single Household											−0.28*	−0.46***		−0.92*
Female Head of HH								0.79**						
Younger Than 25										−1.84*				
Separation	−1.70**													
Work, Education, and Finances														
Same Employer 3rd Q.					−0.71*									
Same Employer 4th Q.											0.33**			
Left Job Prev. Year														−1.04**
In Education								−0.77**						
Vocational Training					−0.85*									
Blue-Collar Worker	−1.16*													
Unemployment Benefits							−1.09***							
Health, Personality, and Activities														
Individual Health Services													−0.71**	
Low Life Satisfaction						−0.60*								
Very Energetic							−0.51*							
Achieved Less													−0.93*	
Vegetarian/Vegan				−0.52*										
Middle-class												0.53*		
Sustainability and Social Justice												−0.56*		

Note: *p <0.05; **p <0.01; ***p <0.001.

In Samples B, C, D, E, and G, fewer than four households were not re-contacted and effects of independent variables could not be reliably identified. Therefore, the weighting factors for these samples in this step are constants and they are omitted from this table.

Table 3.3 – *Continued from previous page*

Explanatory Variable	Sample A	Sample F	Sample H	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1	Sample M2	Sample M3/4	Sample M5	Sample N	Sample O
Outsider Feeling											−0.91***			
Political Party Preference											0.29*			
Many Friends										−1.74*				
No Concerns About Xenophobia											0.24*			
Worried About Own Economic Situation			−0.96*											
Many Single Households												−0.28*		
Building, Area, and Region														
No Balcony/Terrace				−0.44*										
No Garage										2.09**				
No Car in HH											0.29**			
High-Rise Area								−1.07***						
New Apartments												−0.37**		
Manufacturing Sector Low										−1.34**				
Minijobs Low				−0.48*										
Fathers Parental Allowance														−0.89*
Apprentices Low									−0.47*					
Employment Rate High												−0.35**		
Fewer Women Employed				−0.53*										
Commute 300km Or More													−0.93**	
Internal Migration Low									−0.46*					
Development Employment High								−0.78**						
Development Tax Revenue High				−0.79**										
Migration Net Total Low												−0.37**		
Life Expectancy Low												0.37**		
Schleswig-Holstein/Hamburg							−0.66*							
City 20,000 - 100,000 Residents									−0.79**					
High Share of AfD Voters	−0.80*													
Low Share of CDU/CSU Voters											0.43***			
High Share of Green Party Voters													−0.74**	
High Share of Wind Energy								−0.86**						
No. of Observations	1,497	1,947	593	1,797	987	1,086	1,451	621	1,415	580	2,617	1,491	2,335	943
Log Likelihood	-18.36	-27.82	-15.16	-43.11	-18.09	-28.89	-77.05	-41.04	-60.07	-24.74	-372.43	-302.99	-40.21	-32.16

Note: *p <0.05; **p <0.01; ***p <0.001.
In Samples B, C, D, E, and G, fewer than four households were not re-contacted and effects of independent variables could not be reliably identified. Therefore, the weighting factors for these samples in this step are constants and they are omitted from this table.

4 Panel Attrition Due to Refusals

In each panel wave, the second step in successful re-interviewing after having identified the location of households from the preceding wave is to obtain each household's confirmation of willingness to participate in the survey. We define successful re-interviewing relative only to survey-related panel attrition, such as refusals, and ignore survey-unrelated attrition, such as the death of a participant or her decision to move abroad, to generate the longitudinal weights.

4.1 The Frequency of Participation

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

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

Table 4.1: The Frequency of Re-Contacted Households and the Percentage of Participation, Subsamples A to Q 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	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1984	4528		1393																			
1985	4611	89.8	1326	89.1																		
1986	4442	89.2	1290	87.4																		
1987	4194	93.2	1204	92.7																		
1988	4105	91.2	1180	90.8																		
1989	3949	92.4	1146	91.0																		
1990	3871	93.3	1111	92.5	2179																	
1991	3842	94.0	1143	92.4	2213	91.7																
1992	3833	93.5	1144	92.7	2290	88.2																
1993	3838	93.9	1156	92.0	2208	89.2																
1994	3821	93.6	1139	89.8	2122	92.3	236															
1995	3766	93.6	1097	89.5	2101	92.2	540	96.7														
1996	3734	93.3	1061	90.5	2092	93.3	542	91.9														
1997	3674	94.1	1029	90.5	2076	93.5	537	89.2														
1998	3645	92.9	1013	88.6	2066	91.3	523	84.3	1056													
1999	3616	92.0	969	88.5	2030	93.3	495	85.9	1084	81.7												
2000	3535	91.7	929	88.3	2018	93.1	466	91.2	959	87.8	6043											
2001	3448	91.9	899	90.0	2028	91.2	450	88.4	913	88.8	6100	80.5										
2002	3396	92.0	869	88.1	1996	91.1	449	89.5	868	89.1	5420	84.6	1224									
2003	3318	92.6	837	88.6	1974	91.5	432	92.4	828	89.9	4951	88.6	1047	87.0								
2004	3253	92.5	800	89.3	1955	92.7	435	89.2	795	92.1	4719	89.7	1007	89.8								
2005	3214	91.4	774	90.2	1954	90.6	426	89.0	782	90.3	4564	89.2	998	88.1								
2006	3130	90.1	767	85.4	1930	89.0	420	85.7	768	89.3	4370	89.1	990	86.8	1506							
2007	2992	91.0	721	85.2	1832	90.3	385	89.6	725	89.2	4138	89.3	926	89.0	1523	78.0						
2008	2850	90.7	671	84.9	1759	90.5	370	88.6	678	88.8	3939	89.2	901	87.3	1321	81.9						
2009	2723	89.0	616	81.2	1693	90.7	350	87.4	636	90.3	3746	88.2	866	87.4	1142	87.2	1495					
2010	2565	87.5	545	80.9	1627	88.3	333	83.5	604	91.6	3523	86.7	825	90.1	1054	86.6	1709	68.8				
2011	2417	88.9	491	79.6	1538	88.1	301	88.4	589	92.5	3308	87.2	794	88.9	988	86.8			3136			
2012	2285	89.0	439	78.8	1465	89.6	286	87.8	115	80.0	3073	87.9	772	89.0	927	88.2			3179	80.4	1526	
2013	2172	89.7	390	82.3	1413	88.5	267	86.9	98	83.7	2873	89.3	730	92.7	873	89.7			2857	80.7	1549	82.7
2014	2065	90.8	359	84.1	1346	90.0	249	85.5	90	86.7	2732	88.4	720	89.0	823	88.9			2497	84.5	1439	82.5
2015	1986	88.6	329	81.5	1294	87.4	229	84.3	83	84.3	2577	88.2	694	87.3	788	86.8			2296	86.4	1299	85.3
2016	1853	87.9	295	77.3	1213	88.5	207	83.6	80	85.0	2398	87.3	660	89.4	717	89.1			2108	89.3	1201	87.1
2017	1736	88.0	268	75.0	1120	89.0	183	90.2	75	89.3	2199	89.5	619	90.6	673	88.3			2001	88.8	1099	89.8
2018	1631	86.1	236	75.0	1051	88.4	173	85.0	69	85.5	2058	88.0	602	88.5	639	85.8			1902	89.0	1056	88.4
2019	1491	86.0	199	75.9	986	84.2	155	87.7	64	85.9	1940	85.2	573	88.8	589	83.4			1783	86.3	981	85.3

Year	Sample L1		Sample L2		Sample L3		Sample M1		Sample M2		Sample M3/4		Sample M5		Sample N		Sample O		Sample P		Sample Q	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
2010	2074		2500																			
2011	2057	80.1	2228	87.9	924																	
2012	1859	78.9	2222	85.8	931	87.2																
2013	1741	78.2	2151	83.9	912	82.9	2723															
2014	1503	83.0	1990	71.2	824	72.7	2793	72.0														
2015	1396	84.8	1850	74.5	772	76.3	2407	69.3	1096													
2016	1280	87.7	1712	73.9	719	75.0	2067	72.2	1064	62.0	3289											
2017	1200	87.9	1564	79.7	671	77.8	1757	76.8	913	61.2	3138[30]	73.0	1519		2314							
2018	1138	87.1	1465	79.9	638	79.3	1579	76.2	679	71.7	2922	69.7	1477	68.0	2460	85.6	935					
2019	1079	82.9	1418	79.1	604	78.0	1389	74.2	561	69.7	2440	72.3	1316	70.6	2318	82.0	926	67.3	1960		477	

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

4.2 Predicting the Probability of Re-Interviewing vs. Refusal in the Year 2019

Based on the household and interview characteristics measured in the year 2018, and some regional information measured in 2019, we aim at predicting the probability of agreement vs. refusal to participate in the survey for households that were re-contacted in 2019. 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 only use model specifications where all included regressors are to be considered statistically significant (that is different from zero). The definition of the regressors is given in Table 4.2. Table 4.3 reports the subsample-specific estimates of logit models for the probability of participating relative to refusing to participate. Note again that the estimates of regression models of the previous waves 1985 through 2018 are not reported in the present documentation due to space restrictions. These can as well be found in previous attrition reports (e.g. Siegers et al. (2020)).

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

Variable	Label	Value
Interview Characteristics		
New Household	Household new in SOEP	0/1
New Address	Household moved	0/1
Change In HH Composition	Recent change in household composition	0/1
Temp. Drop-Out	Temporary drop-out of household in previous year	0/1
Interview Related HH	Successful interview of related household	0/1
Temp. Drop-Out Related HH	Temporary drop-out of related household	0/1
Drop-Out Related HH	Ultimate drop-out of related household	0/1
Email Known	Email address disclosed	0/1
Phone Unknown	Telephone number undisclosed	0/1
Change Of Interviewer	Latest interview conducted by different interviewer	0/1
CAPI	Computer Assisted Personal Interview	0/1
Relatives Same Interviewer	Relatives were interviewed by the same interviewer	0/1
New Person Two Waves Ago	New person moved into household two waves ago	0/1
Late Interview	Interview was conducted in later months of interviewing period	0/1
Mother-Child-Questionnaire	Additional mother-child-questionnaire in household	0/1
Mode Change	Change of interview mode between the last two waves	0/1
Part. Unit Nonresponse	Household member(s) did not participate last wave	0/1
No Consent Record Linkage	Head of household did not consent to linking SOEP data with other records	0/1
Not Original Sample Member	Head of household is not an original sample member	0/1
Demographic Characteristics		
Family Household	Four or more persons live in household	0/1
Female Head Of HH	Head of household is female	0/1
Single Household	One-person household	0/1
Child Under 12	At least one child under the age of 12 in household	0/1
Younger Than 25	Head of household is younger than 25	0/1
Not Born In Ger	Head of household was not born in Germany	0/1
Foreigner In HH	At least one person who was born outside of Germany in household	0/1
Islamic Countries	Head of household is from an islamic country	0/1
Work, Education, Finances		
Low Education	Head of household has low education (CASMIN 1a - 1c)	0/1
Retired	Head of household is retired	0/1

Table 4.2 – *Continued from previous page*

Variable	Label	Value
Same Employer 3rd Q.	Head of household with current employer since third quarter	0/1
Left Job Prev. Year	Head of household left job at the beginning of the previous year	0/1
In Education	Head of household is currently pursuing education	0/1
Vocational Training	Head of household completed vocational training previous year	0/1
Unemployment Benefits	Head of household currently receives unemployment benefits (ALG II)	0/1
Unemployed Person In HH	At least one unemployed person in household	0/1
Augment Unemp. Benefits	Head of household augments unemployment benefits with paid work	0/1
House Owner	Head of household owns a house	0/1
High Disposable Income	High disposable income	0/1
No Car In HH	No car in household	0/1
No Bank Account In Ger	Head of household does not have a German bank account	0/1
Employed In Public Sector	Head of household is employed in the public sector	0/1
High Income	High income	0/1
BAMF-Integration Course	Head of household took BAMF integration course	0/1
No Asylum Benefits	Household does not receive asylum seeker benefits (AsylbLG)	0/1
Health, Personality, Activities		
Individual Health Services	Head of household received health care not covered by insurance	0/1
Disabled	Head of household has a disability	0/1
Health: Fit	Head of household: health does not affect ability to perform physically demanding tasks	0/1
Vegetarian/Vegan	Head of household does not eat meat	0/1
Middle-Class	Middle-class household (MICROM Sinus-Geo-Milieu)	0/1
Sustainability And Social Justice	Sustainability and social justice are highly valued (MICROM Sinus-Geo-Milieu)	0/1
Frugal Conservative	Frugal conservative values (MICROM Sinus-Geo-Milieu)	0/1
Political Party Preference	Head of household has a general party preference	0/1
Many Friends	Head of household has many close friends	0/1
No Xenophobia Concerns	Head of household has no concerns about xenophobia in the general population	0/1
No Worries Long-Term Peace	Head of household is not worried about long-term peace	0/1
No Economic Concerns	Head of household is not worried about general economic development	0/1
Not Worried About Immigration	Head of household is not worried about immiration into Germany	0/1
Calm And Relaxed	Head of household often feels calm and relaxed	0/1
Chronic Back Pain	Head of household has chronic back pain	0/1
Overweight/High BMI	Head of household is overweight	0/1
Vaping	Head of household vapes/uses e-cigarettes	0/1
Often Eats Poultry	Head of household often eats poultry	0/1

Table 4.2 – *Continued from previous page*

Variable	Label	Value
Thinking About Money: Daily	Head of household thinks about money at least daily	0/1
Social Contacts Abroad	Head of household (refugee) often contacts family or friends in home country	0/1
Unhappy	Head of household rarely felt happy in the past four weeks	0/1
Retirement Concerns	Head of household is concerned about retirement	0/1
Worries About Immigration	Head of household is concerned about immiration into Germany	0/1
Sometimes Feeling Down	Head of household sometimes feels down	0/1
Smoker	Head of household smokes	0/1
Worries Own Economic Situation	Head of household worries about own economic situation	0/1
Achieved Less	Head of household achieved less than desired due to emotional problems	0/1
Often Pressed For Time	Head of household often feels pressed for time	0/1
Did Not Vote	Head of household did not vote in last election	0/1
Not Satisfied w/ Family Life	Head of household is not satisfied with family life	0/1
Donation	Head of household donated to charity last year	0/1
Leisure Time	Head of household spends at least three hours per weekday on leisure activities	0/1
Building, Area, and Region		
Freedom Of Press	Low freedom of press in head of household's home country	0/1
High Amount Of Garbage	Household is located in area that produces high amounts of garbage	0/1
5+ Bedroom Apartments	Household is located in area with high number of 5+ bedroom apartments	0/1
Manufacturing Sector Low	Household is located in area with weak manufacturing sector	0/1
Manufacturing Sector High	Household is located in area with strong manufacturing sector	0/1
Young Unempl. Persons High	Household is located in area with high share of young unemployed persons	0/1
Farm Land	Household is located in area high share of farm land	0/1
Fathers Parental Allowance	High share of fathers receiving parental allowance	0/1
Share Of Water Area High	Household is located in area with high share of water area	0/1
Women Age 65+ Low	Household is located in area with low share of women 65 years and older	0/1
Women Age 65+ High	Household is located in area with high share of women 65 years and older	0/1
Women Age 18-24 Low	Household is located in area with low share of women 18-24 years	0/1
Unemployment Rate Low	Household is located in area with low unemployment rate	0/1
Doctors Per Resident Low	Household is located in area with low number of doctors per resident	0/1
Migration Net Total High	Household is located in area with a high net total of migration	0/1
Building Plot Prices High	Household is located in area with high building plot prices	0/1
Workers w/ College Degree High	Household is located in area with high share of workers with college degree	0/1
Empl. Rate Place of Work High	High employment rate in town/city where head of household works	0/1
Empl. Rate Residence High	High employment rate in town/city where head of household lives	0/1

Table 4.2 – *Continued from previous page*

Variable	Label	Value
Preschoolers Day Care Low	Household is located in area with low share of preschoolers receiving day care	0/1
Population Development Low	Household is located in area with low population development	0/1
Internal Migration Low	Household is located in area with low internal migration	0/1
GDP Per Capita Low	Household is located in area with low GDP per capita	0/1
Service Sector Low	Household is located in area with weak service sector	0/1
Broadband Availability Low	Household is located in area with low broadband availability	0/1
Naturilizations	Household is located in area with high number of naturalizations per immigrant	0/1
Development Employment Low	Household is located in area with weak employment development	0/1
Development Tax Revenue High	Household is located in area with strong tax revenue development	0/1
Development Apartments High	Household is located in area with strong development of new apartments	0/1
Reachability Highways Low	Household is located in area with poor highway reachability	0/1
Intercity Railway Reachable	Household is located in area with good reachability of intercity railway stations	0/1
Residents Under 3 Years Low	Household is located in area with low share of residents younger than 3 years	0/1
Residents Age 30-49 Low	Household is located in area with low share of residents age 30-49	0/1
Residents Age 30-49 High	Household is located in area with high share of residents age 30-49	0/1
Fertility Rate High	Household is located in area with high fertility rate	0/1
Emigration Rate Low	Household is located in area with low emigration rate	0/1
Share Of Women Low	Household is located in area with low share of women	0/1
Share Of Women High	Household is located in area with high share of women	0/1
Women In Local Gov. Low	Household is located in area with low share of women in local government	0/1
Highly Qualified High	Household is located in area with high number of highly qualified workers	0/1
Internet Affinity	Household is located in area with above average internet use	0/1
New Apartment Buildings High	Household is located in area with high number of new apartment buildings	0/1
Local Pharmacies Low	Household is located in area with low number of local pharmacies	0/1
Distance Public Transport High	Household is located in area with high average distance to nearest public transport stop	0/1
Cars Per 1,000 Residents Low	Household is located in area with low number of cars per 1,000 residents	0/1
Cars Per 1,000 Residents High	Household is located in area with high number of cars per 1,000 residents	0/1
Commute 150km And More Low	Household located in area with low number of commutes of 150km and more	0/1
Commute 300km And More High	Household located in area with high number of commutes of 300km and more	0/1
Net Total Commuters Low	Household located in area with similar numbers of in- and out-commuters	0/1
Workers w/ College Degree Low	Household located in area with low share of workers with college degree	0/1
Empl. Rate Women To Men High	Household located in area with high ratio of	0/1
Ratio Young To Old Fit To Work	Household is located in area with high ratio of young to old people who are fit to work	0/1
Refugees Chance	Head of household sees refugees coming to Germany as an opportunity rather than a risk	0/1

Table 4.2 – *Continued from previous page*

Variable	Label	Value
Refugees Positive: Germany	Head of household thinks Germany will become better place for living b/c of refugess	0/1
Refugees Negative: Economy	Head of household thinks refugees will have negative impact on economy	0/1
Saxony	Household is located in Saxony	0/1
Lower Saxony/Bremen	Household is located in Lower Saxony/Bremen	0/1
Berlin/Brandenburg	Household is located in Berlin/Brandenburg	0/1
Elementary Schools Low	Household is located in area with low share of elementary school students	0/1
City Under 20,000 Residents	Household is located in city with fewer than 20,000 residents	0/1
Unskilled Work	Household member(s) job(s) corresponds to semi-/unskilled work	0/1
Fiscal Capacity High	Household is located in area with high fiscal capacity	0/1
Low Share Green Party Voters	Household is located in area with lwo share of Green Party voters	0/1
Open-Ended Contract	Head of household has a permanent employment contract	0/1
Change Building Plot Prices Low	Household is located in area with relatively constant building plot prices	0/1
Change Dwelling Area Low	Household is located in area with relatively constant dwelling area	0/1
Gender Income Gap Low	Household is located in area with low gender income gap	0/1
Investment Securities	Household owned savings or investment securities last year	0/1
Water Area Per Resident Low	Household is located in area with low water area per resident	0/1
Wind Energy Low	Household is located in area with low share of wind energy	0/1
Central Place Class. Cat 0	Central place classification: category 0 no classification	0/1
Mig. Into Town/City High	High number of people moving into town/city where household is located	0/1

Table 4.3: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2019

Explanatory Variable	**Sample A**	**Sample B**	**Sample C**	**Sample D**	**Sample E**	**Sample F**	**Sample G**	**Sample H**	**Sample J**	**Sample K**	**Sample L1**	**Sample L2**	**Sample L3**	**Sample M1**	**Sample M2**	**Sample M3/4**	**Sample M5**	**Sample N**	**Sample O**
Intercept	0.47***	0.18	0.79***	1.70***	1.15***	0.64***	1.27***	1.30***	0.93***	1.18***	0.57***	1.09***	−0.14	0.83***	0.21*	−0.40***	0.51***	0.30	0.97***
Interview Characteristics																			
New Household	−1.37***		−1.01**			−0.86***			−0.77***	−1.00**	−0.62*	−0.67***	−0.77**					−0.51**	
New Address	−0.37*								−0.52***				−0.46*	−0.32*		−0.54***	−0.44***		−0.64**
Change In HH Composition								−0.80**											
Temp. Drop-Out	−1.94***		−1.48***			−1.69***	−1.36***	−1.56***	−1.54***	−1.74***	−0.85***	−1.07***	−1.04***	−1.34***	−1.20***		−0.66***	−1.69***	
Interview Related HH												0.22*	0.60***						
Temp. Drop-Out Related HH			−0.70***																
Drop-Out Related HH																−0.35*			
Email Known													0.40**				0.28*		
Phone Unknown		−1.18**					−0.58**	−1.39***	−0.81***	−0.60*	−0.49*	−1.00***	−1.00***	−0.74***	−0.80***	−1.31***	−1.67***	−0.41***	−0.59***
Change Of Interviewer						−0.37***			−0.25*		−0.39**	−0.33**		−0.63***				−0.15*	
CAPI	0.32***																		
Relatives Same Interviewer			0.44***																
New Person Two Waves Ago																		0.45**	
Late Interview						−0.32***		−0.31*	−0.17*	−0.28**								−0.39***	
Mother-Child-Questionnaire	−0.27*																		
Mode Change						0.31*													
Part. Unit Nonresponse									−0.19*			−0.37***					−0.30**	−0.38***	−0.26*
No Consent Record Linkage																−0.70**			
Not Original Sample Member									−0.37***		−0.31**	−0.36***	−0.29*	−0.32***					
Demographic Characteristics																			
Family Household					−1.30*											0.32***	0.27**	−0.19*	
Female Head Of HH													0.33*				0.20*		
Single Household												−0.28**		−0.23*				−0.26***	
Child Under 12									−0.34***				0.51***						
Younger Than 25																	−0.38***		
Not Born In Ger											−0.30**								
Foreigner In HH														−0.26***					−0.42***
Islamic Countries						0.20*													
No Asylum Benefits																	0.19*		
Open-Ended Contract												−0.21*							
Retired																		0.20*	
Work, Education, Finances																			
Same Employer 3rd Q.		−1.16**																	−0.41**
Left Job Prev. Year				−1.56**									0.47**						
In Education																	0.40*		
Vocational Training										−0.68*									

Note: *p <0.05; **p <0.01; ***p <0.001.

Table 4.3 – *Continued from previous page*

Explanatory Variable	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1	Sample M2	Sample M3/4	Sample M5	Sample N	Sample O
Low Education																			–0.22*
Unemployment Benefits																0.16*			
Unemployed Person In HH																0.21**			
Augment Unemp. Benefits																0.14*			
House Owner		0.56*																	0.42**
High Disposable Income	0.24**												0.41*						
BAMF-Integration Course																	0.22*		
No Car In HH																0.29***			
No Bank Account In Ger																–0.58*			
Employed In Public Sector												0.29**				0.41*			
High Income								0.40**											
Investment Securities									0.38**		0.32*								
Health, Personality, Activities																			
Individual Health Services																			0.24*
Disabled																		0.24*	
Health: Fit						0.14*													–0.26**
Vegetarian/Vegan	–0.49**		–0.52*																
Middle-Class	0.32**																		
Sustainability And Social Justice																	–0.49*		
Frugal Conservative																0.18*			–0.37**
Political Party Preference						0.21***													
Many Friends																	0.46***		0.58***
No Xenophobia Concerns			–0.40***													0.13*			
No Worries Long-Term Peace							–0.54**												
No Economic Concerns											0.22*								
Not Worried About Immigration	0.22**																		
Calm And Relaxed																0.13*			
Chronic Back Pain		0.48*									0.26**								
Overweight/High BMI															0.45**			–0.15*	
Vaping	–0.37*																		
Often Eats Poultry	–0.22*																		
Social Contacts Abroad																0.13*	–0.24*		
Thinking About Money: Daily									0.19*										
Unhappy	0.20**																		
Retirement Concerns									0.18*										
Worries About Immigration											0.21*								
Sometimes Feeling Down													0.30*						
Smoker									0.19*										
Worries Own Economic Situation				–0.97*													–0.21*		
Achieved Less																	0.37*		

Note: *p <0.05; **p <0.01; ***p <0.001.

Table 4.3 – *Continued from previous page*

Explanatory Variable	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1	Sample M2	Sample M3/4	Sample M5	Sample N	Sample O
Often Pressed For Time												0.16*							
Did Not Vote			−0.30*																
Not Satisfied W/ Family Life			−0.26**																
Donation																		0.21***	
Leisure Time																−0.40**			
Freedom Of Press																−0.81*			
Building, Area, Region																			
High Amount Of Garbage									0.21*										
5+ Bedroom Apartments																		0.30***	
Manufacturing Sector Low						0.26***													
Manufacturing Sector High										−0.39***									
Young Unempl. Persons High												−0.25**							
Farm Land																			0.52**
Fathers Parental Allowance														−0.26**					
Share Of Water Area High															0.58***				
Women Age 65+ Low												0.36***							
Women Age 65+ High																		0.21**	
Women Age 18-24 Low		−0.58*						−0.41**											
Unemployment Rate Low												−0.25**							
Doctors Per Resident Low															0.45**				
Migration Net Total High													−0.43**						
Building Plot Prices High															−0.57***				
Workers w/ College Degree High					−0.76*														
Empl. Rate Place of Work High							−0.31*												
Empl. Rate Residence High	−0.20*																		
Preschoolers Day Care Low							0.47**												
Population Development Low		−0.60*																	
Internal Migration Low																			−0.46***
GDP Per Capita Low								0.45**											
Service Sector Low													0.32*						
Broadband Availability Low							−0.47*												
Naturilizations													0.48**						
Development Employment Low								−0.41**											
Development Tax Revenue High								−0.41**											
Development Apartments High											−0.27**					0.31***			
Reachability Highways Low	0.17*																		
Intercity Railway Reachable																0.23***			
Residents Under 3 Years Low																0.27***			
Residents Age 30-49 Low											0.52***								
Residents Age 30-49 High								−0.33*											

Note: *p <0.05; **p <0.01; ***p <0.001.

Table 4.3 – *Continued from previous page*

Explanatory Variable	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H	Sample J	Sample K	Sample L1	Sample L2	Sample L3	Sample M1	Sample M2	Sample M3/4	Sample M5	Sample N	Sample O
Fertility Rate High															−0.47**				
Emigration Rate Low			−0.29**																
Share Of Women Low											0.23*								
Share Of Women High																		0.20*	
Women In Local Gov. Low																	0.21*		
Highly Qualified High							−0.30*												
Internet Affinity																			−0.37**
New Apartment Buildings High																			−0.27*
Local Pharmacies Low	0.32***																		
Distance Public Transport High												−0.26**							
Cars Per 1,000 Residents Low				−1.16**															
Cars Per 1,000 Residents High																−0.36***			
Commute 150km And More Low											−0.40***								−0.84***
Commute 300km And More High																	−0.34***		
Net Total Commuters Low											0.34***								
Workers w/ College Degree Low									−0.20**										
Empl. Rate Women To Men High			0.33**																
Ratio Young To Old Fit To Work														0.18*					
Refugees Chance			−0.35**																
Refugees Positive: Germany									0.36*										
Refugees Negative: Economy														0.25**	0.30*				
Saxony																		0.34**	
Lower Saxony/Bremen												0.45***							
Berlin/Brandenburg																		0.29*	
Elementary Schools Low										−0.27**									
City Under 20,000 Residents									−0.22*					0.21*					
Unskilled Work															0.56**	0.20*			
Fiscal Capacity High						0.19**													
Low Share Green Party Voters											−0.40***								
Change Building Plot Prices Low								−0.36**											
Change Dwelling Area Low										−0.28**									
Gender Income GapLow																0.25***			
Water Area Per Resident Low		0.63**																	
Wind Energy Low																0.13*			
Central Place Class. Cat 0							−0.48*							−0.42**					
Mig. Into Town/City High						−0.17*										−0.18**			
Number of Observations	1,497.00	200.00	988.00	95.00	64.00	1,947.00	574.00	593.00	1,797.00	987.00	1,086.00	1,451.00	621.00	1,415.00	580.00	2,617.00	1,491.00	2,335.00	943.00
Log Likelihood	-498.48	-83.66	-370.52	-25.37	-20.28	-721.80	-161.16	-202.17	-583.39	-348.94	-403.47	-574.20	-235.60	-663.41	-292.94	-1,286.63	-717.39	-921.63	-510.19

Note: *p <0.05; **p <0.01; ***p <0.001.

5 Margins used in the Post-Stratification Process

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

Table 5.1: Marginal Distributions - Household Level

Variables	**Marginal Distributions**
Federal State[31] (*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
Size of Municipality (*Mun. Size*)	Less than 20,000 inhabitants 20,000-100,000 inhabitants 100,000-500,000 inhabitants More than 500,000 inhabitants
Household Size (*H. Type*)	1 ǀ 2 ǀ 3 ǀ 4 ǀ 5 or more members
Houseowner (*Owner*)	Owner ǀ Tenant

[31] 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

Table 5.1 – *Continued from previous page*

Variables	**Marginal Distributions**
Household Typology (*H. Type*)	Single household 2 adults without children 2 adults, 1 or 2 children Single parent, less than 3 children Single parent, 3 or more children Families with more than 3 children Remaining households
Migration Second Generation (*Migr.*)	No 2nd generation migrant in household At least one 2nd gen. migrant in household born after 1995 At least one 2nd gen. migrant in household born 1975-1994 At least one 2nd gen. migr. born 1975-1994 and one after 1995
Nationality (*Nat.*)	EU Country ǀ Former Yugoslavia ǀTurkey ǀ CIS countries Rest of the world ǀ Only German nationality
Year of Immigration (*Imm. Year*)	1900-1979 ǀ 1980-1984 ǀ 1985-1989 ǀ 1990-1994 1995-1999 ǀ 2000-2004 ǀ 2005-2009 ǀ 2010-2013[32]ǀ Other
Target Population AB (*AB*)	Household size and country of origin *(altogether 47 combinations)*
Target Population E, F (*E, F*)	West Germany, all household members German West Germany, at least one household member without Ger. nat. East Germany
Target Population G (*G*)	West Germany, household income <DM 7,500 East Germany, household income <DM 7,500 West Germany, household income DM 7,500-10,000 East Germany, household income DM 7,500-10,000 West Germany, household income >DM 10,000 East Germany, household income >DM 10,000

[32]The additional category "2010-2013" is used from 2015 on

Table 5.1 – *Continued from previous page*

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

Table 5.1 – *Continued from previous page*

Variables	**Marginal Distributions**
Target Population H, J, K (*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)
Household Size and Number of Employed Household Members[33] (*Empl.*)	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 (*ALG*)	Household in West Germany receiving ALG II[34] Household in West Germany without ALG II Household in East Germany receiving ALG II Household in East Germany without ALG II
Greater Regions (*Reg.*)	North Germany ∣ East Germany South Germany ∣ West Germany

[33] *Sample J*: sorted by East and West Germany
[34] Arbeitslosengeld II

Table 5.1 – *Continued from previous page*

Variables	Marginal Distributions
Target Population M1 (*M1*)[35]	1st Generation, 1995-2004, Turkey 1st Generation, 1995-2004, Spain/Greece/Italy 1st Generation, 1995-2004, Poland 1st Generation, 1995-2004, CIS countries 1st Generation, 1995-2004, Arabic Countries 1st Generation, 1995-2004, Late repatriate 1st Generation, 1995-2004, Rest of the world 1st Generation, after 2005, Turkey, Spain, Greece 1st Generation, after 2005, Poland 1st Generation, after 2005, CIS countries 1st Generation, after 2005, Rest of the world 2nd Generation, Not Turkey 2nd Generation, Turkey
Target Population M2 (*M2*)[35]	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
Target Population M3/4 (*M3/4*)	(At least one) M3/4-eligible Person Moved into existing household Household founded by M3/4-eligible person(s) M3/4-eligible Person(s) living in refugee shelter HH not M3/4-eligible

[35]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 5.1 – *Continued from previous page*

Variables	**Marginal Distributions**
Target Population M5 (*M5*)	(At least one) M5-eligible Person Moved into existing household Household founded by M5-eligible person(s) M5-eligible Person(s) living in refugee shelter HH not M5-eligible
Target Population N (*N*)	Part of target population of Sample N Not part of target population of Sample N[36]
Refugee in Household (*Ref.*)	(At least one) Person in household came to Germany as a refugee between 2013 and 2016 No person in household that came to Germany as a refugee between 2013 and 2016[37]
Target Population O (*O*)	Part of "Soziale Stadt"-area Western Germany Part of "Soziale Stadt"-area Eastern Germany HH not part of target population of Sample O
Target Population P (*P*)	bottom wealth tercile, female, young bottom wealth tercile, female, old bottom wealth tercile, male, young bottom wealth tercile, male, old middle wealth tercile, female, young middle wealth tercile, female, old middle wealth tercile, male, young middle wealth tercile, male, old top wealth tercile, female, young top wealth tercile, female, old top wealth tercile, male, young top wealth tercile, male, old
Target Population Q (*Q*)	no lesbian/gay/bisexual person in HH at least one lesbian/gay/bisexual person in HH lesbian/gay/bisexual couple in HH
Target Population P/Q (*P/Q*)	neither P nor Q P, not Q Q, not P P and Q

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

[37]The term "refugee" refers to the target populations of Samples M3/4 in 2016 and Samples M3/4 and M5 from 2017 on.

Table 5.2: Margins - Household Level

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

Table 5.2 – *Continued from previous page*

Year (Samples)	Fed. State	Mun. Size	H. Size	Owner	H. Type	Migr.	Nat.	Imm. Year	Reg.	M1	M2	M3/4	M5	N	Ref.	O	P	Q	P/Q
2013 (A-M1)	+ ∗ M1	+ ∗ M1	+ ∗ M1	+ ∗	+ ∗	+ ∗	+	+	M1	+ M1									
2014 (A-M1)	+	+	+	+	+	+	+	+											
2015 (A-M2)	+ ∗ M2	+ ∗ M2	+ ∗ M2	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗	M2		+ M2								
2016 (A-M3/4)	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗				∗			+				
2017 (A-N)	+ ∗ N	+ ∗ N	+ ∗ N	+ ∗	+ ∗ N	+ ∗ N	+ ∗ N	+ ∗ N					∗	∗	+				
2018 (A-O)	+ ∗ O	+ ∗ O	+ ∗ O	+ ∗	+ ∗ O	+ ∗ O	+ ∗ O	+ ∗ O							+ ∗	∗ O			
2019 (A-Q)	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗							+ ∗		∗	∗	P Q

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

Table 5.3: Marginal Distributions - Person Level

Variables	Distributions
Age and Gender	0-4 male ǀ 0-4 female ǀ 5-9 male ǀ 5-9 female 10-14 male ǀ 10-14 female ǀ 15-19 male ǀ 15-19 female 20-24 male ǀ 20-24 female ǀ 25-29 male ǀ 25-29 female 30-34 male ǀ 30-34 female ǀ 35-39 male ǀ 35-39 female 40-44 male ǀ 40-44 female ǀ 45-49 male ǀ 45-49 female 50-54 male ǀ 50-54 female ǀ 55-59 male ǀ 55-59 female 60-64 male ǀ 60-64 female ǀ 65-69 male ǀ 65-69 female 70+ male ǀ 70+ female
Household Typology (*H. Type*)	1 adult and 0 children ǀ 2 adults and 0 children 3 adults and 0 children ǀ 4 or more adults and 0 children 1 adult and 1 or more children ǀ 2 adults and 1 child 2 adults and 2 children ǀ 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 (*German*)	German nationality ǀ Other nationality
Migration Second Generation (*Migrant 2nd Gen.*)	Indirect migration, born after 1995 Indirect migration, German nat., born 1975/1994 Indirect migration, other nat., born 1975/1994 Indirect migration, other nat. born before 1964 until 1974 Direct or no migration, or indirect migration, but German nationality and born before 1975
Foreign Nationality (*Nation.*)	EU Country ǀ Former Yugoslavia ǀ CIS countries ǀ Turkey Rest of the world ǀ Only German nationality
Year of Immigration (*Imm. Year*)	1900-1979 ǀ 1980-1984 ǀ 1985-1989 ǀ 1990-1994 1995-1999 ǀ 2000-2004 ǀ 2005-2009 ǀ 2010-2013[38] ǀ Other
Target Population G (*G*)	West Germany, household income <DM 7,500 East Germany, household income <DM 7,500 West Germany, household income DM 7,500-10,000 East Germany, household income DM 7,500-10,000 West Germany, household income >DM 10,000 East Germany, household income >DM 10,000

[38] An adjusted category “2010-2013” is used from 2015 on.

Table 5.3 – *Continued from previous page*

Variables	Distributions
Age[39]	0-4 ǀ 5-9 ǀ 10-14 ǀ 15-19 ǀ 20-24 ǀ 25-29 ǀ 30-34 35-39 ǀ 40-44 ǀ 45-49 ǀ 50-54 ǀ 55-59 ǀ 60-64 ǀ 65+
Gender	Male ǀ Female
Target Population L1 (*L1*)	*Four different variables:* Household with child born in 2007 (yes/no) Household with child born in 2008 (yes/no) Household with child born in 2009 (yes/no) Household with child born in the 1st quarter of 2010 (yes/no)
Target Population L2 (*L2*)	Family with low income (LI) Single parent household (SP) Household with at least 3 children (3+) (LI) and (SP) household (LI) and (3+) household (SP) and (3+) household (LI), (SP) and (3+) household
Target Population L3 (*L3*)	Single parent household (SP) Household with at least 3 children (3+) (SP) and (3+) household
Type of Migration Background (*Migrant*)	Immigration before 1995 Immigration between 1995 and 2004 Immigration since 2005 Migration background (indirect) No migration background Not eligible for sample M1

[39] Different categorisation in:
Sample L1: 0, 1, 2, 3, 4-7, 8-12, 13-18, 19-26, 27-31, 32-36, 37-41, 42-46, 47+
Sample L2: 0-3, 4-7, 8-12, 13-18, 19-26, 27-31, 32-36, 37-41, 42-46, 47-51, 52-56, 57+
Sample L3: 0-3, 4-6, 7-11, 12-17, 18-25, 26-30, 31-35, 36-40, 41-45, 46-50, 51-55, 56+
Sample M1: For respondents younger than 19 years old: only one category (0-19).

Table 5.3 – *Continued from previous page*

Variables	Distributions
Target Population M1 (*M1*)	1st generation, earlier than 1995, Turkey, m/f[40] 1st generation, earlier than 1995, Spain/Greece/Italy, m/f 1st generation, earlier than 1995, Late repatriate, m/f 1st generation, earlier than 1995, Rest of the world, m/f 1st generation, 1995-2004, Turkey, m/f 1st generation, 1995-2004, Spain/Greece/Italy, m/f 1st generation, 1995-2004, Poland, m/f 1st generation, 1995-2004, CIS countries, m/f 1st generation, 1995-2004, Arabic countries, m/f 1st generation, 1995-2004, Late repatriate, m/f 1st generation, 1995-2004, Rest of the world, m/f 1st generation, after 2005, Spain/Greece/Italy, m/f 1st generation, after 2005, Poland, m/f 1st generation, after 2005, CIS countries, m/f 1st generation, after 2005, Rest of the world, m/f 2nd generation, Not Turkey, m/f 2nd generation, Turkey, m/f German, m/f Not eligible for sample M1
Target Population M2 (*M2*)	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

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

Table 5.3 – *Continued from previous page*

Variables	Distributions
Part of Target Population of Sample M3/4 (*Ref. M3/4*)	Came to Germany as a refugee between January 2013 and January 2016 Not part of target population of "Refugee Samples"[41]
Federal State - Refugee Samples (*Ref. 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 Not part of target population of "Refugee Samples"
Registered Date of Arrival - Refugee Samples (*Ref. Reg.*)	Arrival including January 2013 to January 2016 Arrival including February 2016 to December 2016 Not part of target population of "Refugee Samples"
Date of Arrival in Germany[42]- Refugee Samples - By Year and Quarter (*Ref. Arrival*)	2013 Q1 ǀ 2013 Q2 ǀ 2013 Q3 ǀ 2013 Q4 2014 Q1 ǀ 2014 Q2 ǀ 2014 Q3 ǀ 2014 Q4 2015 Q1 ǀ 2015 Q2 ǀ 2015 Q3 ǀ 2015 Q4 2016 Q1 ǀ 2016 Q2 ǀ 2016 Q3 ǀ 2016 Q4 Not part of target population of "Refugee Samples"
Country of Origin - Refugee Samples (*Ref. Origin*)	Syria Afghanistan Iraq Albania, Serbia, Kosovo Eritrea, Somalia Iran, Pakistan Other Not part of target population of "Refugee Samples"
Age - Refugee Samples (*Ref. Age*)	0-4 ǀ 5-9 ǀ 10-14 ǀ 15-17 ǀ 18-20 ǀ 21-24 ǀ 25-29 ǀ 30-34 35-39 ǀ 40-44 ǀ 45-49 ǀ 50-54 ǀ 55-59 ǀ 60+ Not part of target population of "Refugee Samples"

[41]The variables marked here with *"Refugee Samples"* refer to the target populations of Samples M3/4 in 2016 and Samples M3/4 and M5 in 2017 respectively.

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

Table 5.3 – *Continued from previous page*

Variables	Distributions
Age and Gender - Refugee Samples (*Ref. Age & Gender*)	0-4 male ǀ 0-4 female ǀ 5-9 male ǀ 5-9 female 10-14 male ǀ 10-14 female ǀ 15-17 male ǀ 15-17 female 18-24 male ǀ 18-24 female ǀ 25-29 male ǀ 25-29 female 30-34 male ǀ 30-34 female ǀ 35-39 male ǀ 35-39 female 40+ male ǀ 40+ female Not part of target population of "Refugee Samples"
Target Population P (*P*)	bottom wealth tercile, female, young bottom wealth tercile, female, old bottom wealth tercile, male, young bottom wealth tercile, male, old middle wealth tercile, female, young middle wealth tercile, female, old middle wealth tercile, male, young middle wealth tercile, male, old top wealth tercile, female, young top wealth tercile, female, old top wealth tercile, male, young top wealth tercile, male, old
Federal States Rural/Urban - Sample P (*P State Urban*)	Schleswig-Holstein, rural Schleswig-Holstein, urban Hamburg Lower Saxony, rural Lower Saxony, urban Bremen North Rhine-Westphalia, rural North Rhine-Westphalia, urban Hesse, rural Hesse, urban Rhineland-Palatinate, rural Rhineland-Palatinate, urban Baden-Württemberg, rural Baden-Württemberg, urban Bavaria, rural Bavaria, urban Saarland Berlin Brandenburg Mecklenburg-West Pomerania Saxony, rural Saxony, urban Saxony-Anhalt Thuringia

Table 5.3 – *Continued from previous page*

Variables	**Distributions**
Age Group - Sample P (*P Age*)	min - 1954 1955 - 1964 1965 - 1969 1970 - 1974 1975 - max

Table 5.4: Margins - Person Level

Year (Samples)	Age & Gender	H. Type	Ger.	Mig. 2nd Gen.	Nat.	Imm. Year	G	Age	Gender	L1	L2	L3	Mig.	M1	M2	Ref. M3/4	Ref. Fed. State	Ref. Reg.	Ref. Arr.	Ref. Orig.	Ref. Age	Ref. Age & Gender	P	P State Urb.	P Age
1984 (A-B)	+	+	+																						
1985 (A-B)	+	+	+																						
1986 (A-B)	+	+	+																						
1987 (A-B)	+	+	+																						
1988 (A-B)	+	+	+																						
1989 (A-B)	+	+	+																						
1990 (A-B)	+	+	+																						
1991 (A-B)	+	+	+																						
1992 (A-B)	+	+	+																						
1993 (A-B)	+	+	+																						
1994 (A-B)	+	+	+																						
1995 (A-B)	+	+	+																						
1996 (A-B)	+	+	+																						
1997 (A-B)	+	+	+																						
1998 (A-E)	+ ∗ E	+ ∗ E	+ ∗ E																						
1999 (A-E)	+	+	+																						
2000 (A-F)	+ ∗ F	+ ∗ F	+ ∗ F																						
2001 (A-F)	+	+	+																						
2002 (A-H)	+ ∗	+ ∗	+ ∗				∗ G																		
2003 (A-H)	+	+	+																						
2004 (A-H)	+	+	+																						
2005 (A-H)	+	+	+																						
2006 (A-H)	+ ∗ H	+ ∗ H	+ ∗ H																						
2007 (A-H)	+	+	+																						
2008 (A-H)	+	+	+																						
2009 (A-I)	+	+	+																						
2010 (A-L2)	+ ∗	+ ∗	+ ∗	+ ∗				L1 L2	L1 L2	L1	L2														
2011 (A-L3)	+ ∗ J	+ ∗ J	+ ∗ J	+ ∗ J				L3	L3			L3													
2012 (A-K)	+ ∗ K	+ ∗ K	+ ∗ K	+ ∗ K																					

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

Table 5.4 – *Continued from previous page*

Year (Samples)	Age & Gender	H. Type	Ger.	Mig. 2nd Gen.	Nat.	Imm. Year	G	Age	Gender	L1	L2	L3	Migrant	M1	M2	Ref. M3/4	Ref. Fed. State	Ref. Reg.	Ref. Arr.	Ref. Orig.	Ref. Age	Ref.Age & Gender	P	P State Urb.	P Age
2013 (A-M1)	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗		M1					M1	∗ M1											
2014 (A-M1)	+	+	+	+	+	+																			
2015 (A-M2)	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗		M2							M2										
2016 (A-M3/4)	+ ∗	+ ∗ M3/4	+ ∗	+ ∗	+ ∗	+ ∗											+ M3/4		+ M3/4	M3/4	+ M3/4	+ M3/4			
2017 (A-N)	+ ∗ N	+ ∗ M5	+ ∗ N	+ ∗ N	+ ∗ N	+ ∗ N										∗	+ M5	+ M5	+ M5	+ M5	+ M5	+ M5			
2018 (A-O)	+ ∗ O	+ ∗	+ ∗ O	+ ∗ O	+ ∗ O	+ ∗ O											+ ∗		+ ∗	+ ∗	+ ∗	+ ∗			
2019 (A-Q)	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗	+ ∗			P								+ ∗		+ ∗	+ ∗	+ ∗	+ ∗	P	P	P

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

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

Based on the regression models of successful vs. unsuccessful re-contacts and agreements vs. refusals to participate, we derive two sets of predicted probabilities, the product of which is the household's "staying probability". The inverse of the probability of staying in the SOEP in 2019 based on characteristics measured in 2018, variable BJHBLEIB, lends itself as a longitudinal weighting variable which itself corrects for selective attrition between waves 2018 and 2019. Tables 6.1, Table 6.2, Table 6.3, Table 6.4, Table 6.5 and Table 6.6 report some subsample specific summary statistics of the longitudinal weights in each wave.

The product of the cross-sectional weight in 2018, variable BIHHRF, and the longitudinal weight in 2019, variable BJHBLEIB, provide the raw data for the cross-sectional weight in 2019. 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 2019.
Tables 6.7 and 6.8 report subsample specific summary statistics of the derived cross-sectional weighting variable BJHHRF and in comparison all previous cross-sectional weights AHHRF through BIHHRF.

Table 6.1: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples A through D (Percentiles of $HBLEIB up to Wave 36).

Year	Sample A				Sample B				Sample C				Sample D			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
1985	1.06	1.10	1.22	4,141	1.09	1.10	1.26	1,181								
1986	1.04	1.07	1.26	3,962	1.10	1.10	1.29	1,128								
1987	1.03	1.03	1.13	3,910	1.03	1.03	1.14	1,116								
1988	1.02	1.04	1.20	3,743	1.03	1.04	1.22	1,071								
1989	1.03	1.04	1.16	3,647	1.03	1.04	1.14	1,043								
1990	1.02	1.02	1.11	3,612	1.04	1.04	1.12	1,028								
1991	1.02	1.02	1.09	3,613	1.03	1.03	1.16	1,056	1.03	1.06	1.18	2,030				
1992	1.01	1.02	1.11	3,585	1.01	1.03	1.16	1,060	1.06	1.06	1.22	2,020				
1993	1.01	1.01	1.16	3,603	1.02	1.03	1.22	1,064	1.03	1.04	1.17	1,970				
1994	1.02	1.02	1.15	3,577	1.03	1.05	1.22	1,023	1.02	1.04	1.12	1,959				
1995	1.01	1.01	1.16	3,526	1.02	1.05	1.29	982	1.03	1.03	1.11	1,938				
1996	1.01	1.03	1.12	3,485	1.04	1.04	1.21	960	1.01	1.02	1.15	1,951	1.00	1.08	1.16	396
1997	1.01	1.02	1.13	3,458	1.02	1.04	1.29	931	1.02	1.04	1.12	1,942	1.05	1.09	1.09	340
1998	1.02	1.03	1.14	3,387	1.04	1.07	1.23	898	1.02	1.02	1.20	1,886	1.08	1.08	1.35	308
1999	1.02	1.02	1.20	3,325	1.04	1.04	1.22	858	1.01	1.03	1.10	1,894	1.05	1.05	1.27	300
2000	1.02	1.02	1.15	3,240	1.03	1.03	1.18	820	1.01	1.03	1.13	1,879	1.02	1.02	1.10	302
2001	1.02	1.02	1.18	3,168	1.02	1.02	1.23	809	1.02	1.02	1.16	1,850	1.03	1.03	1.18	286
2002	1.01	1.02	1.21	3,123	1.04	1.04	1.37	766	1.01	1.02	1.21	1,818	1.00	1.02	1.21	289
2003	1.01	1.03	1.14	3,072	1.01	1.03	1.31	742	1.01	1.03	1.14	1,807	1.01	1.01	1.09	290
2004	1.01	1.01	1.12	3,010	1.04	1.04	1.13	714	1.00	1.01	1.12	1,813	1.00	1.01	1.25	277
2005	1.02	1.02	1.16	2,937	1.05	1.05	1.17	698	1.00	1.02	1.15	1,771	1.00	1.02	1.34	273
2006	1.01	1.04	1.22	2,821	1.01	1.05	1.33	655	1.01	1.04	1.24	1,717	1.03	1.04	1.44	261
2007	1.01	1.03	1.14	2,723	1.03	1.07	1.24	614	1.00	1.03	1.15	1,654	1.01	1.04	1.12	248
2008	1.02	1.05	1.13	2,584	1.01	1.07	1.25	570	1.01	1.03	1.18	1,592	1.02	1.07	1.22	231
2009	1.02	1.05	1.25	2,423	1.01	1.05	1.60	500	1.00	1.03	1.21	1,535	1.00	1.02	1.16	217
2010	1.01	1.06	1.38	2,245	1.01	1.10	1.47	441	1.01	1.04	1.32	1,437	1.00	1.01	1.43	278
2011	1.00	1.04	1.27	2,148	1.01	1.07	1.55	391	1.01	1.05	1.24	1,355	1.01	1.02	1.28	266
2012	1.02	1.08	1.27	2,033	1.01	1.13	1.65	346	1.00	1.05	1.29	1,312	1.00	1.04	1.45	251
2013	1.01	1.06	1.25	1,949	1.01	1.09	1.58	321	1.01	1.07	1.27	1,250	1.01	1.06	1.39	232
2014	1.01	1.04	1.25	1,874	1.01	1.03	1.48	302	1.01	1.04	1.22	1,212	1.00	1.03	1.31	213
2015	1.01	1.06	1.29	1,760	1.01	1.09	1.61	268	1.02	1.07	1.37	1,131	1.00	1.02	1.63	117
2016	1.03	1.08	1.24	1,629	1.01	1.10	1.86	228	1.01	1.07	1.30	1,073	1.01	1.07	1.43	103
2017	1.02	1.09	1.24	1,528	1.02	1.17	1.79	201	1.02	1.08	1.22	997	1.02	1.02	1.28	99
2018	1.03	1.10	1.31	1,404	1.00	1.02	1.98	177	1.03	1.07	1.23	929	1.04	1.04	1.35	92
2019	1.03	1.10	1.28	1,282	1.03	1.15	2.05	151	1.03	1.11	1.40	830	1.00	1.00	1.46	83

Table 6.2: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples E through G (Percentiles of $HBLEIB up to Wave 36).

Year	Sample E				Sample F				Sample G			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
1998												
1999	1.00	1.23	1.47	886								
2000	1.03	1.07	1.21	838								
2001	1.01	1.05	1.25	811	1.08	1.14	1.59	4,911				
2002	1.01	1.02	1.20	773	1.03	1.05	1.46	4,586				
2003	1.04	1.04	1.15	744	1.02	1.04	1.24	4,386	1.06	1.10	1.17	911
2004	1.00	1.01	1.08	732	1.02	1.03	1.19	4,235	1.02	1.03	1.25	904
2005	1.01	1.03	1.18	706	1.01	1.03	1.17	4,070	1.03	1.06	1.25	879
2006	1.00	1.03	1.21	686	1.01	1.03	1.29	3,895	1.00	1.04	1.31	859
2007	1.01	1.01	1.16	647	1.01	1.03	1.15	3,694	1.01	1.05	1.17	824
2008	1.00	1.01	1.19	602	1.01	1.03	1.14	3,513	1.01	1.03	1.18	787
2009	1.00	1.04	1.17	574	1.02	1.04	1.34	3,303	1.02	1.04	1.36	757
2010	1.01	1.04	1.25	553	1.01	1.05	1.40	3,055	1.00	1.01	1.23	743
2011	1.00	1.00	1.17	545	1.01	1.05	1.34	2,885	1.00	1.03	1.35	706
2012	1.05	1.24	1.66	92	1.02	1.08	1.30	2,702	1.02	1.07	1.24	687
2013	1.07	1.20	1.32	82	1.01	1.06	1.21	2,567	1.02	1.05	1.15	677
2014	1.03	1.03	1.42	78	1.02	1.05	1.25	2,414	1.01	1.07	1.32	641
2015	1.13	1.13	1.42	70	1.01	1.05	1.30	2,273	1.01	1.07	1.38	606
2016	1.06	1.06	1.38	68	1.03	1.08	1.24	2,094	1.02	1.02	1.26	590
2017	1.02	1.02	1.45	67	1.03	1.10	1.25	1,968	1.02	1.06	1.22	561
2018	1.03	1.03	1.36	59	1.03	1.08	1.24	1,811	1.02	1.08	1.28	533
2019	1.04	1.04	1.30	55	1.05	1.12	1.32	1,652	1.01	1.08	1.25	509

Table 6.3: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples H, J and K (Percentiles of $HBLEIB up to Wave 36).

Year	Sample H				Sample J				Sample K			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
2007	1.04	1.16	1.46	1,188								
2008	1.01	1.03	1.18	1,082								
2009	1.01	1.03	1.22	996								
2010	1.01	1.04	1.37	913								
2011	1.00	1.05	1.31	858								
2012	1.00	1.03	1.36	818	1.05	1.19	1.52	2,555				
2013	1.00	1.05	1.27	783	1.03	1.13	1.36	2,305	1.04	1.15	1.47	1,281
2014	1.01	1.05	1.27	732	1.03	1.09	1.31	2,110	1.02	1.09	1.34	1,187
2015	1.01	1.09	1.26	684	1.02	1.06	1.25	1,983	1.02	1.05	1.31	1,108
2016	1.01	1.04	1.29	639	1.02	1.06	1.20	1,883	1.02	1.05	1.27	1,046
2017	1.01	1.05	1.35	594	1.06	1.10	1.22	1,776	1.03	1.07	1.20	987
2018	1.01	1.06	1.37	548	1.02	1.06	1.18	1,692	1.03	1.07	1.13	934
2019	1.02	1.09	1.41	491	1.03	1.09	1.31	1,538	1.04	1.09	1.32	837

Table 6.4: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples L1, L2 and L3 (Percentiles of $HBLEIB up to Wave 36).

Year	Sample L1				Sample L2				Sample L3			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
2011	1.10	1.20	1.46	1,647	1.03	1.12	1.37	1,958				
2012	1.04	1.16	1.58	1,467	1.03	1.11	1.35	1,907	1.01	1.10	1.37	806
2013	1.03	1.11	1.59	1,362	1.03	1.09	1.37	1,805	1.02	1.11	1.47	750
2014	1.02	1.11	1.47	1,247	1.10	1.26	1.67	1,416	1.10	1.25	1.76	593
2015	1.01	1.06	1.36	1,184	1.04	1.15	1.91	1,379	1.03	1.12	1.74	582
2016	1.02	1.08	1.25	1,122	1.05	1.16	1.97	1,265	1.03	1.15	1.66	533
2017	1.02	1.06	1.24	1,055	1.03	1.12	1.64	1,247	1.03	1.10	1.93	516
2018	1.03	1.10	1.25	991	1.03	1.15	1.62	1,170	1.05	1.12	1.53	501
2019	1.01	1.13	1.47	894	1.02	1.12	1.75	1,121	1.01	1.11	1.82	466

Table 6.5: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples M1, M2 and M3/M4 (Percentiles of $HBLEIB up to Wave 36).

Year	Sample M1				Sample M2				Sample M3/4			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
2014	1.08	1.28	1.81	2,012								
2015	1.08	1.27	1.89	1,667								
2016	1.07	1.21	1.90	1,493	1.24	1.50	2.23	660				
2017	1.03	1.16	1.61	1,350	1.10	1.36	3.02	559	1.10	1.31	2.06	2,178
2018	1.09	1.18	1.45	1,203	1.04	1.19	2.10	487	1.13	1.37	2.02	2,037
2019	1.07	1.21	1.74	1,030	1.04	1.31	2.00	391	1.06	1.26	2.33	1,763

Table 6.6: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples M5, N, and O (Percentiles of $HBLEIB up to Wave 36).

Year	Sample M5				Sample N				Sample O			
	p10	p50	p90	N	p10	p50	p90	N	p10	p50	p90	N
2018	1.08	1.32	2.37	1,005	1.05	1.13	1.36	2,050				
2019	1.05	1.26	2.65	929	1.03	1.14	1.44	1,889	1.07	1.33	2.20	623

Table 6.7: Summary Statistics of the Derived Cross-Sectional Weights at the Household Level (Percentiles of $HHRF up to Wave 36).

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	3,899	5,081	6,430	8,472	10,033	5,322
1986	539	752	3,602	5,301	6,838	9,280	11,116	5,090
1987	546	790	3,538	5,380	7,043	9,576	11,455	5,026
1988	532	804	3,566	5,638	7,541	10,353	12,536	4,814
1989	552	819	3,598	5,840	7,878	10,810	13,276	4,690
1990	699	1,073	2,217	4,601	7,042	9,897	12,393	6,819
1991	680	1,042	2,329	4,693	7,146	10,289	12,875	6,699
1992	669	1,028	2,337	4,660	7,138	10,533	13,652	6,665
1993	689	1,055	2,403	4,668	7,256	10,753	13,975	6,637
1994	706	1,099	2,403	4,673	7,285	11,208	14,713	6,559
1995	695	1,113	2,386	4,364	6,977	11,081	14,845	6,768
1996	732	1,164	2,392	4,350	7,009	11,388	15,317	6,699
1997	742	1,209	2,400	4,321	7,053	11,854	15,867	6,621
1998	983	1,353	2,331	3,975	6,222	9,884	13,118	7,492
1999	971	1,318	2,307	3,988	6,491	10,879	14,348	7,220
2000	800	1,102	1,760	2,524	3,568	5,083	6,520	13,082
2001	754	1,028	1,755	2,753	4,148	6,096	7,835	11,796
2002	506	657	1,222	2,556	4,191	6,511	8,249	12,320
2003	503	677	1,237	2,563	4,327	6,829	9,082	11,909
2004	491	669	1,213	2,535	4,419	7,261	9,830	11,644
2005	492	678	1,232	2,544	4,517	7,579	10,875	11,294
2006	458	650	1,272	2,393	4,140	6,881	9,762	12,361
2007	458	652	1,255	2,470	4,465	7,585	10,685	11,552
2008	460	656	1,275	2,557	4,752	8,225	11,553	10,921
2009	473	669	1,301	2,630	5,027	9,083	12,454	10,270
2010	220	361	667	1,431	3,658	7,375	11,097	13,888
2011	213	325	611	1,506	3,090	5,580	7,813	16,703
2012	214	325	634	1,641	3,155	5,707	7,608	16,397
2013	179	272	528	1,317	2,971	5,275	7,501	17,992
2014	201	315	625	1,547	3,375	6,118	8,380	15,946
2015	187	299	619	1,491	3,371	6,311	8,900	15,908
2016	41	79	332	1,163	3,109	6,052	8,736	17,715
2017	41	74	325	1,156	2,817	5,446	7,945	19,628
2018	44	93	369	1,244	2,963	5,645	8,382	18,622
2019	48	94	306	1,065	2,860	5,708	8,531	18,971

Table 6.8: Summary Statistics of the Derived Cross-Sectional Weights at the Person Level (Percentiles of $PHRF up to Wave 36).

Year	p5	p10	p25	p50	p75	p90	p95	N
1984	397	553	1,172	4,366	5,226	6,054	6,803	16,173
1985	454	634	1,421	4,620	5,718	6,893	8,065	14,508
1986	488	678	1,533	4,678	6,023	7,591	9,010	13,804
1987	509	719	1,598	4,728	6,224	7,893	9,407	13,563
1988	489	686	1,623	4,895	6,565	8,501	10,210	12,872
1989	528	755	1,746	5,017	6,893	8,964	10,691	12,443
1990	681	1,023	1,900	3,441	6,146	8,286	10,234	18,254
1991	731	1,072	1,914	3,707	6,191	8,477	10,604	17,844
1992	778	1,139	2,000	3,740	6,300	8,724	11,108	17,429
1993	845	1,235	2,088	3,832	6,378	8,999	11,385	17,072
1994	873	1,284	2,110	3,840	6,417	9,273	12,064	16,715
1995	765	1,147	2,018	3,611	6,098	9,065	12,160	17,345
1996	799	1,191	2,027	3,641	6,142	9,414	12,774	16,944
1997	839	1,214	2,064	3,669	6,251	9,690	13,351	16,583
1998	911	1,267	2,041	3,511	5,594	8,507	11,404	18,249
1999	904	1,242	2,015	3,496	5,794	9,238	12,657	17,501
2000	725	973	1,565	2,314	3,215	4,560	5,853	30,784
2001	690	934	1,533	2,452	3,647	5,411	6,932	27,956
2002	445	612	1,064	2,192	3,713	5,818	7,629	29,101
2003	440	620	1,086	2,203	3,810	6,119	8,224	27,867
2004	436	617	1,086	2,186	3,892	6,509	8,845	26,918
2005	441	630	1,118	2,236	4,015	6,879	9,597	25,638
2006	414	597	1,116	2,176	3,681	6,300	8,808	27,442
2007	415	600	1,124	2,230	3,925	6,950	10,096	25,505
2008	425	611	1,155	2,297	4,134	7,651	11,241	23,792
2009	439	627	1,182	2,386	4,358	8,347	12,572	22,096
2010	176	277	534	1,031	2,514	5,405	8,576	35,945
2011	164	251	455	989	2,388	4,409	6,622	42,031
2012	165	246	466	1,109	2,535	4,483	6,710	40,351
2013	140	211	408	907	2,267	4,278	6,258	44,633
2014	156	242	480	1,092	2,602	4,918	7,187	38,839
2015	143	228	471	1,102	2,613	5,071	7,492	38,224
2016	25	41	187	787	2,319	4,804	7,356	44,042
2017	25	41	164	785	2,175	4,376	6,589	48,249
2018	27	47	207	857	2,346	4,670	7,135	44,576
2019	29	52	193	787	2,346	4,941	7,560	43,443

References

AAPOR (2016). *Standard Definitions. Final Dispositions of Case Codes and Outcome Rates for Surveys*. Tech. rep. American Association of Public Opinion.

De Vries, L., M. Fischer, M. Kroh, S. Kühne, and D. Richter (2021). *Design, Nonresponse, and Weighting in the 2019 Sample Q (Queer) of the Socio-Economic Panel.* SOEP Survey Papers 940: SOEP Survey Papers Series C – Data Documentation, DIW/SOEP, Berlin 2021.

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

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

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

— (2011b). *SOEP 1990/91 – Methodenbericht Ostdeutschland zu den Befragungsjahren 1990-1991 (Welle 1/2 – Ost) des Sozio-oekonomischen Panels.* SOEP Survey Papers 14, DIW/SOEP, Berlin 2011.

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

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

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

— (2011f). *SOEP 2000 – Methodenbericht erste Welle der SOEP Stichprobe F zum Befragungsjahr 2000 (Welle 17) des Sozio-oekonomischen Panels.* SOEP Survey Papers 37, DIW/SOEP, Berlin 2011.

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

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

Jacobsen, J., M. Kroh, S. Kühne, J. A. Scheible, R. Siegers, and M. Siegert (2019). *Supplementary of the IAB-BAMF-SOEP Survey of Refugees in Germany (M5) 2017.* SOEP Survey Papers 605, DIW/SOEP, Berlin 2019.

Kara, S., S. Zimmermann, and SOEP-Group (2018). *SOEPcompanion (v34), Release 2018, v.1* SOEP Survey Papers 588, DIW/SOEP, Berlin 2018.

Kühne, S. and M. Kroh (2017). *The 2015 IAB-SOEP Migration Study M2: Sampling Design, Nonresponse, and Weighting Adjustment.* SOEP Survey Papers 473, DIW/SOEP, Berlin 2017.

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

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

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., 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 (2015b). *Gewichtung und Integration von Auffrischungsstichproben am Beispiel des Sozio-oekonomischen Panels (SOEP).* In: Nonresponse Bias: Qualitätssicherung sozialwissenschaftlicher Umfragen. Ed. by J. Schupp and C. Wolf. Wiesbaden: Springer. pp.409–444.

Martin, S., A. Zabal, and B. Rammstedt (2018). "PIAAC-L data collection 2016: technical report." In: *GESIS Papers* 2018|05.

OECD (2016). *Technical Report of the Survey of Adult Skills (PIAAC), 2nd Edtition*. Tech. rep. OECD - Organisation for Economic Co-operation and Development.

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

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

Rendtel, U. (1995). *Lebenslagen im Wandel: Panelausfälle 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., M. Kroh, N. Watson, et al. (2013). "The implementation of cross-sectional weights in household panel surveys". In: *Statistics Surveys* 7, pp. 37–57.

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.

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, C., C. Bartels, K. Göbler, M. M. Grabka, J. König, R. Siegers, and S. Zinn (2020). *Improving the Coverage of the Top-Wealth Population in the Socio-Economic Panel (SOEP).* SOEP Papers on Multidisciplinary Panel Data Research 1114, DIW/SOEP, Berlin 2020.

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.

Siegers, R., V. Belcheva, and T. Silbermann (2020). *Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2018).* SOEP Survey Papers 826, DIW/SOEP, Berlin 2020.

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.

Steinhauer, H. W., M. Kroh, and J. Goebel (2020). *Sampling, Nonresponse, and Weighting in the Sample O.* SOEP Survey Papers 827, DIW/SOEP, Berlin 2020.

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 Ergän-zungsstichprobe H zum Befragungsjahr 2006 (Welle 23) des Sozio-oekonomischen Panels.* SOEP Survey Papers 57, DIW/SOEP, Berlin 2011.

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

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

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

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

Zabal, A., S. Martin, N. Massing, D. Ackermann, S. Helmschrott, I. Barkow, and B. Rammstedt (2014). *PIAAC Germany 2012. Technical Report*. Tech. rep. GESIS - Leibniz Institute for Social Sciences.

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

— (2017). "PIAAC-L data collection 2015: technical report." In: *GESIS Papers* 2017|29.