

Rathje, Martin; Glemser, Axel

**Research Report**

## SOEP-Core - 2020: Report of survey methodology and fieldwork

SOEP Survey Papers, No. 1050

**Provided in Cooperation with:**

German Institute for Economic Research (DIW Berlin)

*Suggested Citation:* Rathje, Martin; Glemser, Axel (2021) : SOEP-Core - 2020: Report of survey methodology and fieldwork, SOEP Survey Papers, No. 1050, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/242106>

**Standard-Nutzungsbedingungen:**

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

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

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

**Terms of use:**

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

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

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



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

# 1050<sup>2021</sup>

**SOEP** Survey Papers  
Series B - Survey Reports (Methodenberichte)

## SOEP-Core – 2020: Report of Survey Methodology and Fieldwork

Martin Rathje, Axel Glemser (Kantar Public)

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>

**Editors:**

Dr. Jan Goebel, DIW Berlin  
Prof. Dr. Stefan Liebig, DIW Berlin and Freie Universität Berlin  
Prof. Dr. David Richter, DIW Berlin and Freie Universität Berlin  
Prof. Dr. Carsten Schröder, DIW Berlin and Freie Universität Berlin  
Prof. Dr. Jürgen Schupp, DIW Berlin and Freie Universität Berlin  
Prof. Dr. Sabine Zinn, DIW Berlin and Humboldt Universität zu Berlin

Please cite this paper as follows:

Martin Rathje, Axel Glemser (Kantar Public). 2021. SOEP-Core – 2020: Report of Survey Methodology and Fieldwork. SOEP Survey Papers 1050: Series B. Berlin: DIW/SOEP.



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.  
© 2021 by SOEP

ISSN: 2193-5580 (online)

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

Contact: [soepapers@diw.de](mailto:soepapers@diw.de)

# **SOEP-Core – 2020: Report of Survey Methodology and Fieldwork**

**Martin Rathje, Axel Glemser (Kantar Public)**

*Munich, 2021*

# KANTAR PUBLIC

## SOEP 2020

## Wave Report+

### SOEP-Core

to

DIW Berlin  
Mohrenstraße 58  
10177 Berlin

Submitted by:  
Martin Rathje

Kantar GmbH  
Kantar Public Deutschland  
Landsberger Str. 284  
80687 München

Your contact persons:

Martin Rathje  
Senior Consultant  
Tel: 089 5600 - 1417  
E-Mail: martin.rathje@kantar.com

Axel Glemser  
Senior Director  
Tel: 089 5600 - 1486  
E-Mail: axel.glemser@kantar.com

Munich, 28 June 2021

316301618, 316401053, 316301743, 316301741, 316301744, 316301745, 316301742

# Contents

<b>Introduction</b>	<b>6</b>
<b>Overview of the SOEP-Core subsamples</b>	<b>8</b>
<b>1 Samples A-Q</b>	<b>11</b>
1.1 Introduction	11
1.2 Development samples A-Q	12
1.3 Structure of the gross sample	15
1.4 Questionnaires and fieldwork material	20
1.4.1 Questionnaires	20
1.4.2 Fieldwork material	23
1.5 Conducting the survey	27
1.5.1 Survey mode	27
1.5.2 Fieldwork timing	30
1.5.3 Translations	31
1.5.4 Panel maintenance and incentives	32
1.5.5 Movers and tracing	33
1.5.6 Interviewer characteristics, training & monitoring	34
1.6 Fieldwork results	40
1.6.1 Participation on household level	40
1.6.2 Participation on individual level	47
1.6.3 Participation by types of questionnaires	51
1.6.4 Interview length per questionnaire	54
1.6.1 Consent to record linkage	55
1.7 Data preparation	56
1.8 Delivered data	58
<b>2 Samples M1/2</b>	<b>59</b>
2.1 Introduction	59
2.2 Development samples M1/2	60

2.3	Structure of the gross sample	62
2.4	Questionnaires and fieldwork material	65
2.4.1	Questionnaires	65
2.4.2	Fieldwork material	66
2.5	Conducting the survey	68
2.5.1	Survey mode	68
2.5.2	Fieldwork timing	70
2.5.3	Translations	70
2.5.4	Panel maintenance and incentives	71
2.5.5	Movers and tracing	73
2.5.6	Interviewer characteristics, training & monitoring	73
2.6	Fieldwork results	77
2.6.1	Participation on household level	77
2.6.2	Participation on individual level	81
2.6.3	Participation by types of questionnaires	84
2.6.4	Interview length per questionnaire	87
2.6.5	Consent to record linkage	88
2.7	Data preparation	88
2.8	Delivered data	89
3	Samples M3-5	90
3.1	Introduction	90
3.2	Development samples M3-5	91
3.3	Structure of the gross sample	93
3.4	Questionnaires and fieldwork material	96
3.4.1	Questionnaires	96
3.4.2	Fieldwork material	97
3.5	Conducting the survey	100
3.5.1	Survey mode	100
3.5.2	Fieldwork timing	100
3.5.3	Translations	100
3.5.4	Panel maintenance and incentives	103
3.5.5	Movers and tracing	104
3.5.6	Interviewer characteristics, training & monitoring	105
3.6	Fieldwork results	108

# KANTAR PUBLIC

3.6.1	Participation on household level	108
3.6.2	Participation on individual level	112
3.6.3	Participation by types of questionnaires	115
3.6.4	Interview length per questionnaire	116
3.6.5	Consent to record linkage	117
3.7	Data preparation	118
3.8	Delivered data	119
4	Sample M6	120
4.1	Introduction	120
4.2	Background Sample M6	121
4.3	Structure of the Gross Sample	121
4.4	Questionnaires and fieldwork material	124
4.4.1	Questionnaires	124
4.4.2	Fieldwork material	125
4.5	Conducting the survey	128
4.5.1	Survey mode	128
4.5.2	Fieldwork timings	128
4.5.3	Translations	129
4.5.4	Panel Maintenance and Incentives	131
4.5.5	Movers and Tracing	131
4.5.6	Interviewer Characteristics, Training & Monitoring	133
4.6	Fieldwork results	135
4.6.1	Participation on household level	135
4.6.2	Participation on individual level	137
4.6.3	Participation by types of questionnaires	138
4.6.4	Interview length per questionnaire	139
4.6.5	Consent to Record Linkage	139
4.7	Data preparation	140
4.8	Delivered data	141
5	Sample M7 / M8a	142
5.1	Introduction	142
5.2	Background Sample M7 / M8a	143
5.3	Structure of the Gross Sample	144
5.4	Questionnaires and fieldwork material	146



5.4.1	Questionnaires	146
5.4.2	Fieldwork material	147
5.5	Conducting the survey	149
5.5.1	Survey mode	149
5.5.2	Fieldwork timings	149
5.5.3	Translations	150
5.5.4	Panel Maintenance and Incentives	152
5.5.5	Movers and Tracing	152
5.5.6	Interviewer Characteristics, Training & Monitoring	154
5.6	Fieldwork results	156
5.6.1	Participation on household level	156
5.6.2	Participation on individual level	158
5.6.3	Participation by types of questionnaires	159
5.6.4	Interview length per questionnaire	160
5.6.5	Consent to Record Linkage	160
5.7	Data preparation	161
5.8	Delivered data	161

## Introduction

The Socio-Economic Panel (SOEP) has been conducted since 1984 as a panel survey to observe the living situation of private households in Germany. In the 2020 survey year, Kantar conducted the 37<sup>th</sup> wave of the SOEP. The SOEP, based at the German Institute for Economic Research (DIW Berlin), is the longest-running and largest multidisciplinary survey in Germany. Data are collected annually on various topics such as living situation, income, employment, education, health, personality traits and attitudes. The SOEP department at the DIW has overall scientific responsibility for the project, which includes distributing the data to researchers worldwide for use in their own studies.

The SOEP can be divided into three main (types of) surveys, only the first of which is addressed in this report:

1. **SOEP-Core** is the centerpiece of the Socio-Economic Panel study launched in 1984. The data provide information on every member of every household taking part in the survey. Respondents include Germans, foreign nationals residing in Germany, recent immigrants, and new samples of refugees added from 2016 onwards.
2. The longitudinal SOEP Innovation Sample (**SOEP-IS**) was created in 2012 as a special sample for testing highly innovative research projects that involve too great a risk of non-response to be included in the long-term SOEP study. Proposals approved for inclusion in SOEP-IS to date include economic behavioral experiments, implicit association tests (IAT), and complex procedures for measuring time use (day reconstruction method, DRM). Researchers at universities and research institutes worldwide are encouraged to submit innovative proposals to SOEP-IS.
3. There are several studies that have incorporated questions from the SOEP questionnaire to validate their results on a representative sample of the German population ("SOEP as Reference Data"). These SOEP-Related Studies (**SOEP-RS**) are designed and implemented in close cooperation with the SOEP team and structured similarly to the SOEP. Another type of SOEP-Related Studies provides participants from SOEP-Core with additional questionnaires or interviews their employers or daycare providers via related sampling. Some examples of SOEP-Related Studies include BASE-II (Berlin Aging Study II), FiD (Families in Germany), PIAAC-L (Programme for International Assessment of Adult Competencies-Longitudinal), SOEP-ECEC Quality (Early Childhood Education and Care Quality in the SOEP), SOEP-LEE (SOEP Employer-Employee Survey), BIP (Bonn Intervention Panel), and BRISE (Bremen Initiative to Foster Early Childhood Development).

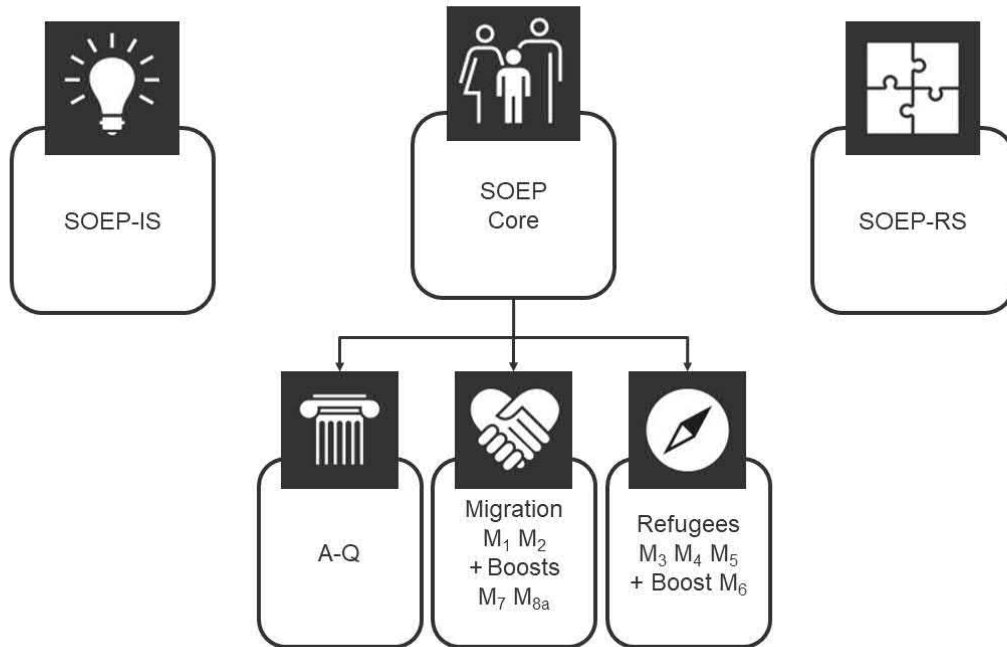
Since the first wave in 1984, Kantar in Munich has been responsible for the implementation of the survey – in particular, the development of the survey instruments, fieldwork, panel maintenance, data checking and processing. The tasks of Kantar also include methodological documentation and reporting. The present report refers to the survey conducted in 2020 in SOEP-Core, which includes the longitudinal samples A-Q (**Part 1**), the migration and refugee samples M1/2 and M3-5 (**Parts 2 and 3**) as well as the most recent boost samples M6 (**Part 4**) and M7 and M8a (**Part 5**) that were established in 2020.

# Overview of the SOEP-Core subsamples

The different SOEP-Core subsamples constitute the centerpiece of the Socio-Economic Panel study. In 2020, it consisted of four different compositions of samples that will be addressed in separate parts in this Wave Report (**Figure 0.1**).

1. Within SOEP-Core, **samples A-Q** form the heart of the SOEP. They contain the oldest samples, beginning with SOEP founding sample A from 1984 and the highest number of participating households (13,460 in 2020). Fieldwork traditionally starts at the beginning of February, and its questionnaires serve as a master for the other SOEP-Core subsamples.
2. The **SOEP migration survey** with its samples M1 and M2 was established in 2013 and is designed to improve the representation of migrants living in Germany. Fieldwork started in April, using the questionnaires from samples A-Q, supplemented by translated questionnaires for five different languages.
3. In order to map recent migration and integration dynamics, **SOEP refugee samples M3 to M5** were installed beginning in the year 2016. In 2020, fieldwork began in August with a questionnaire that was tailored to issues of recent refugees while containing many questions from the SOEP samples A-Q as well.
4. **Sample M6 – a boost sample of refugees** targeted the same population as the older refugee sample M5 - adult refugees who have applied for asylum in Germany since 1 January 2013 and are currently living in Germany – and the same sample design and sample frame were used.
5. The two boost samples, **samples M7 and M8a**, were added the SOEP migration sample system. Like the older migration samples M1 and M2, the Integrated Employment Biographies Sample (IEBS) of the Federal Employment Agency (BA) served as the sampling frame for both boost samples. Boost sample M7's goal was to capture migration dynamics and processes from 2016 to 2018 with a focus on EU migration. To ensure that statistically significant group comparisons can be made, sampling was restricted to the three most significant countries of origin in that time period: Romania, Bulgaria, and Poland. M8a, on the other hand, was designed to help evaluate the skilled worker immigration law (Fachkräfteeinwanderungsgesetz), which came into effect March 1, 2020, and targeted migrants from third countries that came to Germany between 2017 and 2018, sampling them as a control group for a treatment group that will be sampled at a later date.

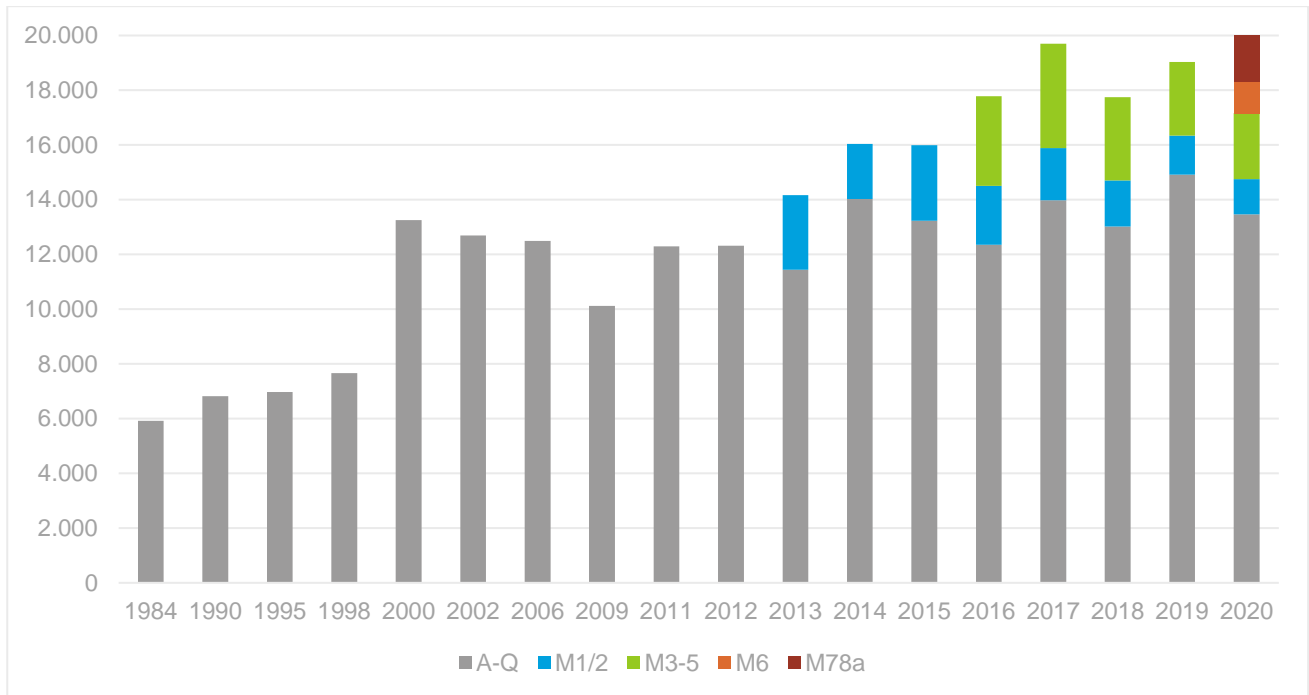
Figure 0.1: Overview of SOEP-Core subsamples



1

Starting with the first 5,924 participating households in 1984, SOEP-Core grew to 20,184 households in 2020. Over the years, more and more samples were added – sometimes samples of all households in Germany, sometimes special populations such as migrants or households with high incomes. In the last couple of years, boosts of the overall sample mainly originated from the migration and refugee samples as well as special boost samples such as samples P and Q (**Figure 0.2**).

**Figure 0.2: Development of SOEP-Core since 1984 – number of households**



The five different groups of samples displayed in **Figure 0.2** will be addressed in five different parts of this report. In the first chapter, fieldwork in the samples A-Q in 2020 will be described in detail, followed by two parts that deal with migration samples M1/2 and M3-5. The last two chapters address samples M6 and M7/8a, the newest additions to SOEP-Core in 2020.

# 1 Samples A-Q

## 1.1 Introduction

**Table 1.1** gives a short overview of the main characteristics of the 2020 wave for the samples A-Q. 13,460 households from samples A-Q participated between February and August. This results in a response rate of 82.2 percent. The adjusted response rate of households that participated in the previous wave was 86.2 percent. In all households, 21,614 adults and 1,270 youths gave interviews. For an additional 1,420 children of various age groups, data are available from interviews with a parent. 14 different questionnaires were fielded in up to five different modes. Partial unit non-response (PUNR: the share of households with more than one household member with at least one missing individual questionnaire) was at 31.0 percent, a higher rate than in the previous wave (26.9 percent). However, this increase was mostly driven by sample N that continued having a relatively high PUNR and by the integration of sample P and Q that have participated for the second time in the survey.

Table 1.1: **Summary fieldwork A-Q**

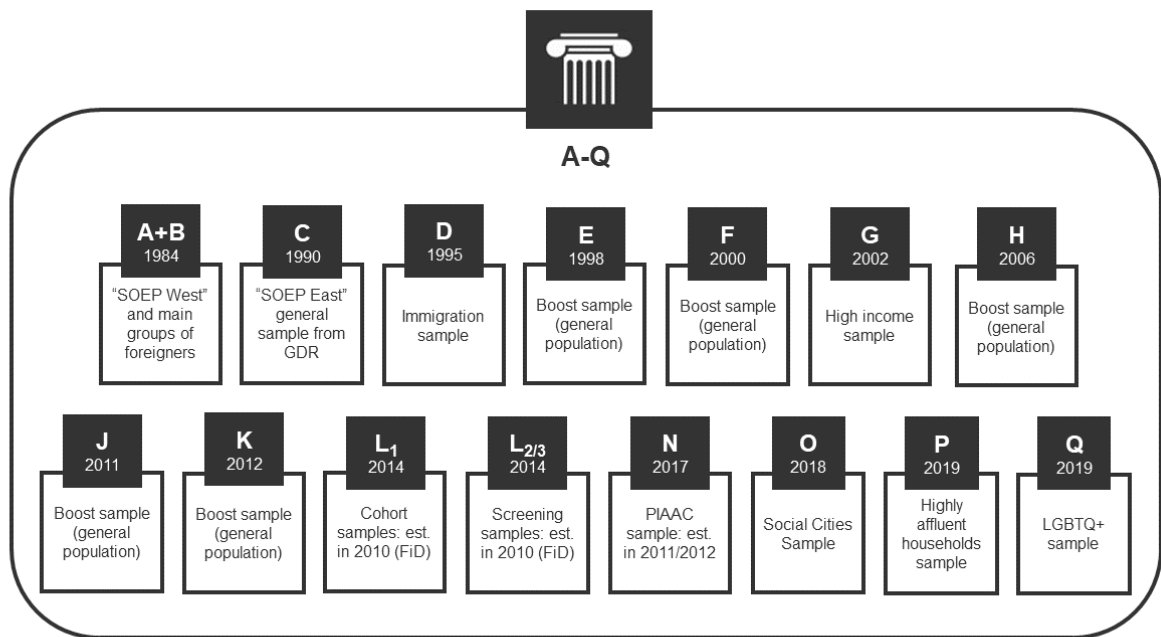
Fieldwork period	February – December
Mode (main questionnaires)	CAPI, PAPI, SELF, MAIL, CAWI
Gross sample (hh)	16,550
Net sample (hh)	13,460
Response rate (adjusted; hh) <sup>1</sup>	Overall: 82.2 Prev. wave respondents: 86.2 Prev. wave dropouts: 41.8 New households: 51.3
Number of questionnaires	Adults: 5 Youths: 4 Children: 5
Net sample (individuals)	Adults: 21,614 Youths: 1,270 Children: 1,420
Questionnaire length (median, in minutes)	Household: 15 Adult: 40
Partial unit non-response	31.0

<sup>1</sup> RR = percentage of all households with at least one hh and individual interview in the gross sample (gross sample adjusted for households where the last person is deceased or the household moved abroad, is permanently untraceable or dissolved households where the last member moved into another SOEP household).

## 1.2 Development samples A-Q

Samples A-Q contain the oldest SOEP samples A and B from 1984 and many other samples that were integrated into the study over the years. Some of them were designed to represent the general population while others were meant to improve sample sizes for special groups such as migrants, households with high income, or families (**Figure 1.1**).

Figure 1.1: Overview of samples A-Q



2



Households and individuals with the longest history of (continuous) panel participation took part for the 37<sup>th</sup> time in 2020 (samples A and B). Another 13 samples were added to SOEP samples A and B since 1984<sup>1</sup>:

- **Sample A “Residents in the Federal Republic of Germany” (1984)** is one of the two initial samples of the SOEP and covers 4,528 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).
- **Sample B “Foreigners in the Federal Republic of Germany” (1984)** is one of the two initial samples of the SOEP and covers 1,393 private households with a Turkish, Greek, Yugoslavian, Spanish, or Italian household head.
- **Sample C “German residents in the German Democratic Republic (GDR)” (1990)** covers persons in 2,179 private households in which the household head was a citizen of the GDR.
- **Sample D “Immigrants” (1994/95)** covers 531 private households in which at least one household member had moved from abroad to West Germany after 1984. It mainly consists of ethnic Germans migrating from Eastern Europe to Germany. This sample includes two subsamples which were drawn independently in 1994 (D1) and in 1995 (D2).
- **Sample E “Refreshment I” (1998)** added another 1,056 households to the SOEP. It is the first sample that was designed to be representative for all private households in both East and West Germany and 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.
- **Sample F “Refreshment II” (2000)** covers 6,043 households and substantially increases the sample size of the SOEP. Households with at least one adult who does not have German citizenship were oversampled in Sample F.
- **Sample G “High income” (2002)** covers 1,224 households with a monthly income of at least 3,835 euros 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 4,500 euros were interviewed further.
- **Sample H “Refreshment III” (2006)** served as a general population refresher covering 1,506 private households in Germany.
- **Sample L1 “Cohort sample” (2010<sup>2</sup>)** covers 2,074 private households in Germany, in which at least one household member is a child born between January 2007 and March 2010. Again, migrants identified by an “onomastic procedure” are oversampled.
- **Samples L2/3 “Family types I/II” (2010<sup>3</sup>)** covers 3,424 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. Since the eligible subpopulation is relatively small and an adequate sampling frame is lacking, a preceding telephone screening procedure identified eligible households.
- **Sample J “Refreshment IV” (2011)** is another general population refresher covering 3,136 households. A disproportional sampling design was implemented in order to increase the number of migrant households in the SOEP.
- **Sample K “Refreshment V” (2012)** is the last general population refresher so far, totaling 1,526 households.
- **Sample N “PIACC sample” (2017)** integrated 2,314 households of former participants of the Programme for the International Assessment of Adult Competencies (PIAAC).
- **Sample O “Social City Sample” (2018)** added 935 households and was designed to enhance the potential of the data for analysis by incorporating more city-specific environments.

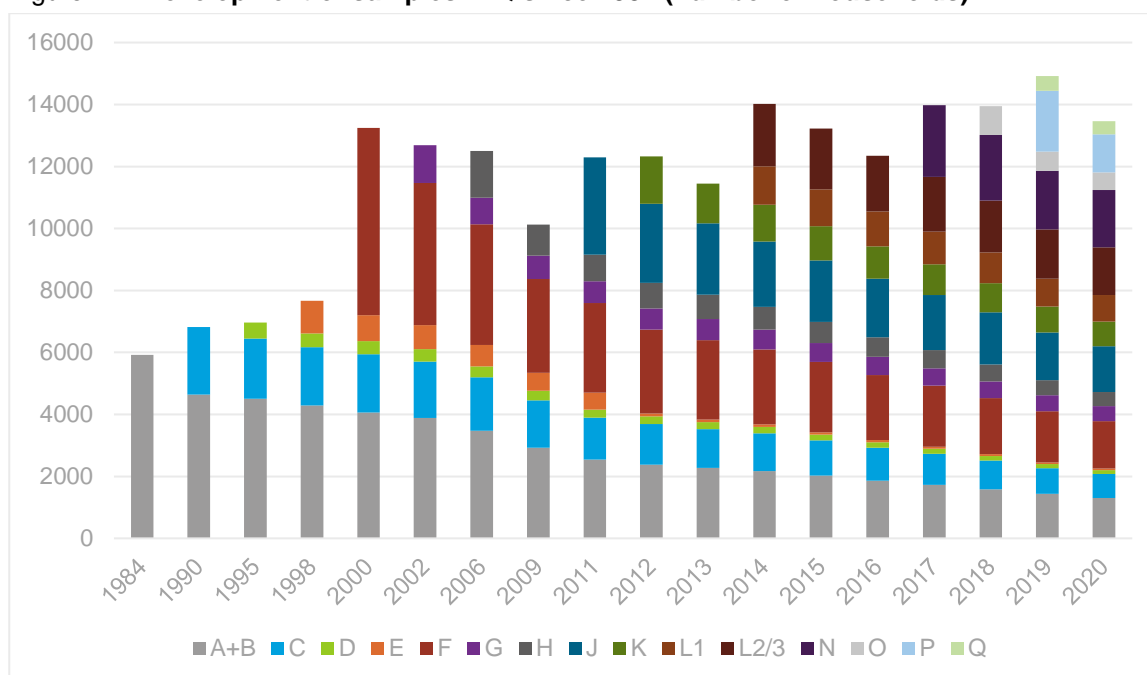
---

<sup>1</sup> Source for sample sizes and descriptions until the year 2012: Martin Kroh, Simon Kühne, Rainer Siegers, Veronika Belcheva. 2018. SOEP-Core – Documentation of Sample Sizes and Panel Attrition (1984 until 2016). SOEP Survey Papers 480: Series C. Berlin: DIW/SOEP.

<sup>2</sup> Samples L1-3 were part of the SOEP-related study “Familien in Deutschland” (FiD) that was established in 2010. After the project ended, the remaining families were integrated into the SOEP (2014).

- **Sample P “Highly affluent households” (2019)** included 1,960 highly affluent households in Germany to compensate a lack of data on wealthy population groups. Due to the hard-to-survey characteristics of the target group, the sample was constructed by using public information on the ownership structures of businesses in the German trade register as people in the top percentile of the wealth distribution are very likely to have some form of equity or shares in a company.
- **Sample Q “LGBTQ+” (2019)** added 477 households of the hard-to-survey population subgroup of lesbians, gays, bisexuals, transgender people, and those who identify as non-binary. Various screening techniques were implemented to arrive at the desired gross sample.

Figure 1.2: Development of samples A-Q since 1984 (number of households)



Since 1984, the SOEP has been extended multiple times by different subsamples, leading to a diverse sampling structure (**Figure 1.2**). Of 5,921 households which comprised the first SOEP subsamples A and B in 1984, 1,305 households remained in the survey in 2020<sup>3</sup>. The total number of households participating in the survey as part of samples A-Q has more than doubled since 1984, reaching a net sample of 13,460 households in 2020.

<sup>3</sup> Including split-off households.

## 1.3 Structure of the gross sample

The result from the previous year's wave forms the basis for the gross sample of the following year (**Table 1.2**). All respondents (i.e., with and without refusal for the next wave) as well as temporary dropouts are transferred into the next wave. Reasons for a temporary dropout can be, e.g., sickness or a lack of time. Households that are not willing or able to take part in future waves are coded as final dropouts and will not receive an invitation for the next year's survey wave. The same applies to households which dropped out temporarily in two consecutive waves ("two-year rule"). In 2020, 14,918 households were coded as respondents and 10,421 as temporary dropouts. Households coded as temporary dropouts in the first wave of sample P and Q were excluded from fieldwork in 2020. Thus, 16,056 households from 2019 form the base sample for the survey wave in 2020.

Accordingly, the administered gross sample in 2020 can be divided into different types of households (**Table 1.3**). Apart from respondents and temporary dropouts of the previous wave, there is a certain number of new households every year which are added to the gross sample. These households are split off from existing households within the survey, e.g., when young adults decide to leave the parental home to form their own household. The administered gross sample in 2020 includes a total of 16,550 households. 90.2 percent of those households also participated in the previous wave, while 6.8 percent had dropped out temporarily. 3.0 percent of the gross sample are new households. The proportions vary between the different groups of subsamples, e.g., in sample O 20.1 percent of respondents were temporary dropouts in previous waves.

Key household characteristics of the different samples are presented in **Table 1.4, Table 1.5, and Table 1.6**. In most samples, the majority of households consists of one or two household members. Samples L1 and L2/3 form an exception here because they originated from the "Families in Germany (FiD)" study that samples different kinds of family types. Further, **Table 1.5** includes an overview of the sample distribution in terms of the different states. In **Table 1.6**, two indicators to classify the household's living environment are presented. The community type (BIK) indicator separates different municipality sizes by the number of inhabitants, distinguishing between center and periphery of city regions. The second community size indicator classifies municipality size into seven categories.

# KANTAR PUBLIC

Table 1.2: Final gross sample 2019 by sample

Household Level	Total		A-H		J-K		L1		L2/3		N		O		P		Q	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample</b>	<b>39,433</b>	<b>100.0</b>	<b>6,063</b>	<b>100.0</b>	<b>2,817</b>	<b>100.0</b>	<b>1,088</b>	<b>100.0</b>	<b>2,079</b>	<b>100.0</b>	<b>2,340</b>	<b>100.0</b>	<b>952</b>	<b>100.0</b>	<b>23,259</b>	<b>100.0</b>	<b>835</b>	<b>100.0</b>
Participating households without refusal for next wave	14,800	37.5	5,015	82.7	2,367	84.0	893	82.1	1,589	76.4	1,881	80.4	625	65.7	1,953	8.4	477	57.1
Participating households with refusal for next wave	118	0.3	91	1.5	8	0.3	1	0.1	3	0.1	8	0.3	0	0.0	7	0.0	0	0.0
Temporary dropouts	10,421	26.4	294	4.8	161	5.7	97	8.9	225	10.8	200	8.5	160	16.8	9,103	39.1	181	21.7
Final dropouts	13,754	34.9	532	8.8	236	8.4	79	7.3	175	8.4	193	8.2	166	17.4	12,196	52.4	177	21.2
2-year rule	340	0.9	131	2.2	45	1.6	18	1.7	87	4.2	58	2.5	1	0.1	0	0.0	0	0.0

Table 1.3: Administered gross sample 2020 by sample

Household Level	Total		A-H		J-K		L1		L2/3		N		O		P		Q	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample</b>	<b>16,550</b>	<b>100.0</b>	<b>5,529</b>	<b>100.0</b>	<b>2,599</b>	<b>100.0</b>	<b>1,029</b>	<b>100.0</b>	<b>1,976</b>	<b>100.0</b>	<b>2,155</b>	<b>100.0</b>	<b>797</b>	<b>100.0</b>	<b>1,983</b>	<b>100.0</b>	<b>482</b>	<b>100.0</b>
Participating households in previous wave	14,924	90.2	5,111	92.4	2,376	91.4	894	86.9	1,592	80.6	1,889	87.7	625	78.4	1,960	98.8	477	99.0
Temp. dropout in previous wave <sup>1</sup>	1,132	6.8	289	5.2	161	6.2	97	9.4	225	11.4	200	9.3	160	20.1	0	0.0	0	0.0
New households (Split-off hh)	494	3.0	129	2.3	62	2.4	38	3.7	159	8.0	66	3.1	12	1.5	23	1.2	5	1.0

<sup>1</sup> The gross sample consists of more cases than anticipated at the end of the previous wave because four households that did not participate in 2019 re-entered the panel. Temporary dropouts from samples P and Q are excluded from gross sample in 2020.

# KANTAR PUBLIC

Table 1.4: Household characteristics by samples I

Household Level	Total		A-H		J-K		L1		L2/3		N		O		P		Q	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample<sup>1</sup></b>	<b>25,179</b>	<b>100.0</b>	<b>5,400</b>	<b>100.0</b>	<b>2,536</b>	<b>100.0</b>	<b>991</b>	<b>100.0</b>	<b>1,817</b>	<b>100.0</b>	<b>2,089</b>	<b>100.0</b>	<b>625</b>	<b>100.0</b>	<b>11,063</b>	<b>100.0</b>	<b>658</b>	<b>100.0</b>
<b>HH size</b>																		
0 unknown	9,284	36.9	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	9,103	82.3	181	27.5
1	4,372	17.4	1,759	32.6	887	35.0	69	7.0	382	21.0	515	24.7	259	41.4	250	2.3	251	38.1
2	5,977	23.7	2,348	43.5	1,044	41.2	63	6.4	491	27.0	819	39.2	202	32.3	855	7.7	155	23.6
3	2,244	8.9	644	11.9	303	11.9	176	17.8	284	15.6	356	17.0	74	11.8	363	3.3	44	6.7
4	2,151	8.5	508	9.4	196	7.7	411	41.5	301	16.6	308	14.7	60	9.6	347	3.1	20	3.0
5+	1,151	4.6	141	2.6	106	4.2	272	27.4	359	19.8	91	4.4	30	4.8	145	1.3	7	1.1

<sup>1</sup> Status as reported at the end of wave 2019. New households and households that re-joined the panel in 2020 but were not part of fieldwork in 2019 are thus missing.

Table 1.5: Household characteristics by samples II

Household Level	Total		A-H		J-K		L1		L2/3		N		O		P		Q	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample<sup>1</sup></b>	<b>25,179</b>	<b>100.0</b>	<b>5,400</b>	<b>100.0</b>	<b>2,536</b>	<b>100.0</b>	<b>991</b>	<b>100.0</b>	<b>1,817</b>	<b>100.0</b>	<b>2,089</b>	<b>100.0</b>	<b>625</b>	<b>100.0</b>	<b>11,063</b>	<b>100.0</b>	<b>658</b>	<b>100.0</b>
<b>State</b>																		
Schleswig-Holstein	746	3.0	149	2.8	112	4.4	57	5.8	97	5.3	81	3.9	16	2.6	213	1.9	21	3.6
Hamburg	584	2.3	95	1.8	46	1.8	11	1.1	29	1.6	45	2.2	35	5.6	296	2.7	27	4.6
Lower Saxony	1,888	7.5	482	8.9	265	10.4	90	9.1	205	11.3	253	12.1	24	3.8	527	4.8	42	7.1
Bremen	163	0.6	35	0.6	11	0.4	9	0.9	21	1.2	25	1.2	10	1.6	46	0.4	6	1.0
North Rhine-Westphalia	4,858	19.3	1,052	19.5	487	19.2	212	21.4	367	20.2	427	20.4	114	18.2	2,088	18.9	111	18.8
Hesse	1,703	6.8	355	6.6	171	6.7	59	6.0	123	6.8	128	6.1	39	6.2	773	7.0	55	9.3
Rhineland Palatinate	1,202	4.8	273	5.1	113	4.5	46	4.6	67	3.7	65	3.1	53	8.5	561	5.1	24	4.1
Baden-Wuerttemberg	2,916	11.6	585	10.8	257	10.1	132	13.3	187	10.3	212	10.1	25	4.0	1,456	13.2	62	10.5
Bavaria	4,311	17.2	784	14.5	433	17.1	165	16.6	292	16.1	330	15.8	79	12.6	2,143	19.4	85	14.4
Saarland	241	1.0	49	0.9	32	1.3	4	0.4	11	0.6	15	0.7	8	1.3	119	1.1	3	0.5
Berlin	1,096	4.4	248	4.6	112	4.4	33	3.3	61	3.4	70	3.4	60	9.6	436	3.9	76	12.9
Brandenburg	1,155	4.6	277	5.1	99	3.9	26	2.6	74	4.1	89	4.3	49	7.8	524	4.7	17	2.9
Mecklenburg Western Pomerania	687	2.7	149	2.8	56	2.2	25	2.5	45	2.5	42	2.0	28	4.5	330	3.0	12	2.0
Saxony	1,712	6.8	420	7.8	163	6.4	60	6.1	103	5.7	157	7.5	22	3.5	760	6.9	27	4.6
Saxony-Anhalt	1,003	4.0	222	4.1	89	3.5	36	3.6	61	3.4	74	3.5	44	7.0	464	4.2	13	2.2
Thuringia	846	3.4	225	4.2	90	3.5	26	2.6	74	4.1	76	3.6	19	3.0	327	3.0	9	1.5

<sup>1</sup> Status as reported at the end of wave 2019. New households and households that re-joined the panel in 2020 but were not part of fieldwork in 2019 are thus missing.

# KANTAR PUBLIC

Table 1.6: Household characteristics by samples III

Household Level	Total		A-H		J-K		L1		L2/3		N		O		P		Q	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample<sup>1</sup></b>	<b>25,271</b>	<b>100.0</b>	<b>5,400</b>	<b>100.0</b>	<b>2,536</b>	<b>100.0</b>	<b>991</b>	<b>100.0</b>	<b>1,817</b>	<b>100.0</b>	<b>2,089</b>	<b>100.0</b>	<b>785</b>	<b>100.0</b>	<b>11,063</b>	<b>100.0</b>	<b>590</b>	<b>100.0</b>
<b>BIK type<sup>2</sup></b>																		
0	6,987	27.6	1,484	27.5	625	24.6	250	25.2	417	22.9	508	24.3	356	45.4	3,082	27.9	265	44.9
1	2,756	10.9	561	10.4	219	8.6	82	8.3	137	7.5	217	10.4	16	2.0	1,483	13.4	41	6.9
2	3,478	13.8	828	15.3	381	15.0	125	12.6	306	16.8	287	13.7	230	29.3	1,224	11.1	97	16.4
3	3,698	14.6	717	13.3	366	14.4	186	18.8	287	15.8	332	15.9	44	5.6	1,710	15.5	56	9.5
4	561	2.2	131	2.4	51	2.0	29	2.9	35	1.9	39	1.9	24	3.1	240	2.2	12	2.0
5	1,691	6.7	405	7.5	177	7.0	58	5.9	146	8.0	199	9.5	64	8.2	616	5.6	26	4.4
6	2,861	11.3	589	10.9	295	11.6	137	13.8	220	12.1	220	10.5	49	6.2	1,305	11.8	46	7.8
7	2,090	8.3	456	8.4	260	10.3	85	8.6	170	9.4	198	9.5	2	0.3	892	8.1	27	4.6
8	670	2.7	135	2.5	107	4.2	20	2.0	66	3.6	61	2.9	0	0.0	272	2.5	9	1.5
9	479	1.9	94	1.7	55	2.2	19	1.9	33	1.8	28	1.3	0	0.0	239	2.2	11	1.9
<b>Community size<sup>3</sup></b>																		
1	1,413	5.6	330	6.1	135	5.3	46	4.6	106	5.8	105	5.0	0	0.0	666	6.0	25	4.2
2	2,330	9.2	436	8.1	246	9.7	114	11.5	169	9.3	232	11.1	0	0.0	1,104	10.0	29	4.9
3	6,683	26.4	1,425	26.4	664	26.2	311	31.4	500	27.5	555	26.6	44	5.6	3,091	27.9	93	15.8
4	4,769	18.9	986	18.3	482	19.0	154	15.5	307	16.9	384	18.4	117	14.9	2,253	20.4	86	14.6
5	1,925	7.6	465	8.6	234	9.2	97	9.8	185	10.2	201	9.6	103	13.1	593	5.4	47	8.0
6	3,765	14.9	846	15.7	347	13.7	98	9.9	296	16.3	290	13.9	267	34.0	1,512	13.7	109	18.5
7	4,386	17.4	912	16.9	428	16.9	171	17.3	254	14.0	322	15.4	254	32.4	1,844	16.7	201	34.1

<sup>1</sup> Status as reported at the end of wave 2019. New households and households that re-joined the panel in 2020 but were not part of fieldwork in 2019 are thus missing.

<sup>2</sup> BIK type: 0 (more than 500,000 inhabitants/center) 1 (more than 500,000 inh./periphery), 2 (100,000 to 499,999 inh./center)

<sup>3</sup> 3 (100,000 to 499,999 inh./periphery), 4 (50,000 to 99,999 inh./center), 5 (50,000 to 99,999 inh./periphery), 6 (20,000 to 49,999 inh.), 7 (5,000 to 19,999 inh.), 8 (2,000 to 4,999 inh.), 9 (fewer than 2,000 inh.).

<sup>3</sup> Community size: 1 (fewer than 2,000 inhabitants), 2 (2,000 to 5,000 inh.), 3 (5,000 to 20,000 inh.), 4 (20,000 to 50,000 inh.), 5 (50,000 to 100,000 inh.), 6 (100,000 to 500,000 inh.), 7 (more than 500,000 inh.).

## 1.4 Questionnaires and fieldwork material

### 1.4.1 Questionnaires

In 2020, a total of 14 questionnaires and instruments were used for the SOEP survey in samples A-Q (**Table 1.7**). All questionnaires, especially the questionnaire for households and the individual questionnaire, are modified and extended on an annual basis. Slight adjustments of the item order or question text usually serve the purpose of methodological advancement. Other modifications are needed in order to comply with societal and political modernization processes to continuously enable adequate measurement. Additionally, several special modules are integrated into the questionnaires following a fixed rotation over multiple years. The following section briefly describes all questionnaires.

Table 1.7: **Questionnaires and modes**

	A-L1, N, O, P, Q	L2/3
	CAPI/PAPI	CAPI/CAWI
Household questionnaire	✓	✓
Individual questionnaire	✓	✓
Life-history questionnaire	✓	✓
Youth questionnaire: age 16 or 17	✓	✓
Cognitive competency tests <sup>1</sup>	✓	-
Youth questionnaire: age 13 or 14	✓	✓
Youth questionnaire: age 11 or 12	✓	✓
Mother and child questionnaire: age 9 or 10	✓	✓
Questionnaire for parents: age 7 or 8	✓	✓
Mother and child questionnaire: age 5 or 6	✓	✓
Mother and child questionnaire: age 2 or 3	✓	✓
Mother and child questionnaire: newborn	✓	✓
Questionnaire "Gap"	✓	✓
Questionnaire "Deceased person"	✓	✓

<sup>1</sup> Only in PAPI mode.



## Household questionnaire

The household questionnaire needs to be completed by only one individual per household, most suitably by the one with best knowledge about the respective matter. It includes questions about:

- Housing, related expenditures, expenditure burden, energy consumption, and mobility
- Household income, social benefits, rental income, and investment returns
- Potential loan burden and opportunity to accumulate assets
- Care-dependent household members, their degree of need and caregivers
- Children living in the household, their schools and care facilities

In the household questionnaire 2020, the biennially asked questions about house equipment like balconies, elevators, etc. were added again. Various questions about the consumption and supply of different forms of energy were rotated into the questionnaire from previous waves. Further, household mobility was assessed with a new question about car sharing services and rotated questions about sustainable forms of mobility. Some new questions regarding municipal benefits and food banks were implemented. Finally, in the children's activities module, respondents were asked to provide some information about children's activities in and outside of school or kindergarten.

## Individual questionnaire

The individual questionnaire is answered by all adult household members (2020: participants born in 2002 or earlier). Usually, more changes and modifications are applied to this questionnaire from year to year than to any of the others. Certain questions and topics are included, following fixed rotations of 2 to 10 years. New items can be developed, or old items reintegrated. The instrument includes questions about:

- Current life situation
- Important changes during the last year
- Current job
- Health and illness
- Attitudes and opinions
- Family situation

In 2020, new questions on an array of different topics were added:

- Employment subject to social insurance contributions
- In-depth assessment about the indicated gross hourly wage
- Digitalization at work and work-life balance
- Jobs, ordering, and renting via internet
- Dual citizenship, identification, and mother language

Also, rotating modules with a fixed replication cycle and questions with irregular appearance were implemented in the individual questionnaire:

- General attitudes towards life, in social interaction, and regarding the personal future
- Daily routine
- Home office
- Timekeeping at work
- Changes in work life, commuting, and professional training opportunities
- Overtime and special bonuses
- Reduction in earning capacity
- Health and disability
- Private and public health insurance
- Blood donation

- Migration and refugees
- Integration indicators

## Additional questionnaires

Apart from the household and individual questionnaires, 12 additional questionnaires were used in 2020. These questionnaires were not answered by all individuals or households in the sample but by particular groups of persons:

- The **life-history questionnaire** is completed by all new respondents joining a panel household (born 2002 or earlier). It mainly collects biographical background information and is applied only once per person. In the previous wave in 2019, it was restructured to better capture the individual migration background.
- The **youth questionnaire age 16 or 17** is answered by household members of the respective age (2020: born in 2003). From 2019 onward it does not replace the individual and the life-history questionnaire anymore resulting in a shorter questionnaire focusing on topics such as school and work, leisure, health, self-perception as well as friendship and family life. It is surveyed once per person.
- In addition to the youth questionnaire: age 16 or 17, the **cognitive competency test** is used after the questionnaire in order to measure the adolescent's level of development. The test entails three exercises with fixed time limits for each one. It is surveyed on paper and only in interviewer-assisted modes since interviewer presence is a necessary prerequisite.
- In 2016, the **youth questionnaire age 13 or 14** was included for the first time to close the gap between the first questioning of young panel members at the age of 11 or 12 and their regular participation in the survey at the age of 19. In 2020 all panel members born in 2006 were asked to complete this questionnaire.
- With the **youth questionnaire age 11 or 12** for household members born in 2008, which was integrated for the first time in 2014, the entry age for young panel members was considerably reduced. The instrument focuses on topics such as school, leisure, health, self-perception as well as friendship and family life.
- For households with children, five more instruments were used in 2020. Mothers of newborn children (born in 2020 or 2019 after the previous survey wave) were asked to complete the **mother and child questionnaire newborn**. Mothers of 2- or 3-year-old children (2020: born in 2017) were asked to answer a respective **mother and child questionnaire 2 or 3**. Similar instruments were available for three more age groups as shown in **Table 1.7**. For children born in 2012, not only the mother but both parents were asked to complete the **questionnaire for parents age 7 or 8**.
- Individuals who temporarily dropped out in the previous wave were asked to complete the **questionnaire "gap"** which is a very short version of the last wave's individual questionnaire. It aims to minimize gaps in longitudinal data on panel members.
- The **questionnaire "deceased person"** is addressed to individuals who lost one of their family members in the year of the survey or the previous year.

The primary interviewing method in samples A-Q is face-to-face with computer-assisted personal interviewing (CAPI) or paper-and-pencil interviewing (PAPI), while a small number of households is interviewed via self-administered mail questionnaires (MAIL). For sample L2/3, a multi-mode design using computer-assisted web interviews (CAWI) as well as regular face-to-face interviewing (CAPI) is applied. As depicted in **Table 1.7**, questionnaires that require an interviewer's presence, such as the cognitive competency test are excluded for this subsample.

## 1.4.2 Fieldwork material

In addition to the questionnaires, a whole range of fieldwork materials, such as letters, leaflets or documents for the interviewers, were designed, printed and sent to households and interviewers. **Table 1.8** provides an overview of the different material types that were prepared in samples A-Q for the different survey modes.

### **Advance letter**

About two weeks before the start of the fieldwork period, households that were to take part in CAPI or PAPI mode received an advance letter in which the interviewer's visit was announced. In samples A-H, which generally do not receive a cash incentive<sup>4</sup>, an announcement letter was not sent to the household head but to every individual household member who took part in the study in the previous wave. This letter enclosed a 5-euro lottery ticket as an unconditional incentive.

For households in the MAIL mode, there were a two-way approach. In households that usually agree to a phone call prior to the actual fieldwork period to update the information about the household structure, all respondents from the previous wave received the 5-euro lottery ticket as an unconditional incentive.<sup>5</sup> After the phone call, these households received another envelope with all the questionnaires they had to fill out in the current wave. Households that had opted out of the phone call but remained in the study had received their questionnaires right away.

Households in CAWI mode received an advance letter with the login information for all household members that had to fill in a questionnaire online.

### **Thank-you letter**

Shortly after participating in the study, each household in modes CAPI, PAPI and MAIL received a thank-you letter with a postage stamp enclosed as a small gift. Respondents from samples A-H (CAPI, PAPI and MAIL mode) that had not received the 5-euro lottery ticket as an unconditional incentive before the start of fieldwork – either because they were new to the panel or because they had not taken part in the previous wave – received an additional, individual thank-you letter with the 5-euro lottery ticket.

Thank-you letters for CAWI respondents also included their incentive in the form of a voucher. The participants could choose at the end of the questionnaire whether they prefer to receive this voucher via e-mail or letter.

---

<sup>4</sup> A small number of households in sample A-H (CAPI, PAPI, MAIL) receives a cash incentive because they used to be part of an incentive experiment (see chapter 1.5.4).

Table 1.8: Overview of fieldwork material

	CAPI	PAPI	MAIL	CAWI
Advance letter	<ul style="list-style-type: none"> <li>• A-H (lottery ticket/cash)</li> <li>• J-K, L1, N, O, Q (cash)</li> <li>• L2-3 (cash)</li> <li>• P (lottery ticket)</li> </ul>	<ul style="list-style-type: none"> <li>• A-H (lottery ticket/cash)</li> </ul>	<ul style="list-style-type: none"> <li>• With short phone interview (lottery ticket/check)</li> <li>• Without short phone interview (lottery ticket/check)</li> </ul>	With login information
Thank-you letter	Thank-you letter	Thank-you letter	Thank-you letter	Thank-you letter including incentive <ul style="list-style-type: none"> <li>• Mail</li> <li>• E-mail</li> </ul>
Reminder letter	-	-	<ul style="list-style-type: none"> <li>• Reminder 1</li> <li>• Reminder 2</li> <li>• Reminder 3</li> </ul>	Reminder 1 (non-response on hh level) Reminder 2 (partial unit non-response)
Special letters during field time	-	-	<ul style="list-style-type: none"> <li>• New households</li> <li>• New person in hh</li> <li>• Mode switchers</li> </ul>	New households New person in hh Mode switchers
Leaflet	<ul style="list-style-type: none"> <li>• Lottery ticket</li> <li>• Cash</li> </ul>	<ul style="list-style-type: none"> <li>• Lottery ticket</li> <li>• Cash</li> </ul>	<ul style="list-style-type: none"> <li>• No cash</li> <li>• Check</li> </ul>	CAWI
Declaration on data protection	<ul style="list-style-type: none"> <li>• A-H, J-K, L1, N, O, P, Q</li> <li>• L2-3</li> </ul>	A-H	A-H	L2-3
Consent to record linkage form	<ul style="list-style-type: none"> <li>• DRV<sup>1</sup> (except P)</li> </ul>	DRV <sup>1</sup>	--	DRV <sup>1</sup>
Address form and household grid <sup>2</sup>	Electronic form (Mein Kantar)	Paper form	Paper form	Electronic form (database mask for phone interviewers)
Showcards	<ul style="list-style-type: none"> <li>• Same for all samples</li> </ul>	--	--	--
Project instruction book	<ul style="list-style-type: none"> <li>• A-H, J-K, L1, N, O, P, Q</li> <li>• L2-3</li> </ul>	A-H	Short instruction manual for phone interviewers updating the hh grid	Handout for phone interviewers updating the hh grid
Other interviewer material	<ul style="list-style-type: none"> <li>• Project description</li> <li>• Contact card</li> <li>• HH information card</li> </ul>	<ul style="list-style-type: none"> <li>• Project description</li> <li>• Contact card</li> <li>• HH information card</li> </ul>	--	--

<sup>1</sup> German statutory pension insurance (Deutsche Rentenversicherung – DRV)

<sup>2</sup> Including the so-called “B3 Form” that is used for processing address changes.

## Reminder letter

In the MAIL and CAWI modes, letters were not only sent out at the beginning and end of the fieldwork period but also in between to remind households to take part in the study. Households in the MAIL mode received up to three reminders. In CAWI mode, two reminders were sent out – one to address non-response on the household level and another one to address missing individual questionnaires.

## Special letters during field time

In MAIL and CAWI mode, there was also a need to send out other, more specialized letters, depending on the situation of the household. If a new household or person was identified during field time, e.g. because a child moved out or a new person joined the household, this new household or person received paper questionnaires or login information for the online questionnaires. Another occasion for special letters during field time were mode switches from CAPI/PAPI to MAIL/CAWI.

## **Leaflet**

Every household received an eight-page leaflet with SOEP reports and published results. In CAPI and PAPI modes, the leaflet was handed over by the interviewer, while it was sent via mail in the MAIL and CAWI modes. There were four slightly different versions, differing by incentive types: cash/lottery ticket (A-Q without L2/3) and cash/voucher (L2/3).

## **Declaration on data protection**

Every household got access to a two-page declaration on data protection detailing the organizations that are responsible for processing all respondent data along with a description of data handling and data recipients. The sheet was handed out by the interviewers in CAPI and PAPI modes and sent via e-mail in the MAIL and CAWI modes. There were two versions of the declaration on data protection because the list of suppliers differs slightly between samples A-Q without L2/3 and L2/3.

## **Consent to record linkage form**

To all respondents born in 2002 or earlier, a consent to record linkage form was presented that allows an individual linkage between the respondent's data and data from the German statutory pension insurance (Deutsche Rentenversicherung – DRV). Members of the CAPI sample could sign either on the touch screen of the interviewer's laptop or on paper. Members of the PAPI and CAWI sample received a personalized consent form on paper. The households participating in MAIL mode did not receive this form. Respondents in sample P and Q were not asked their consent to a record linkage.

## **Address form and household grid**

The address form provided an overview of the household composition as it was last known to Kantar. The interviewers had to document when and with whom the survey was conducted or why a sample member did not participate in the current year. They were also asked to note every single contact attempt made.

Moreover, the interviewers were asked to carefully document any moves of households and household members or changes in the household composition. An additional form had to be completed for every person that had left the household since the last survey because they remain potential respondents and were tracked even if the new address was unknown.

Interviewers did this electronically in the "Mein Kantar" software for CAPI and in paper form for PAPI. For the households in the MAIL mode that received a phone call prior to the start of fieldwork, the contact person who called them filled out a paper form. For participants in MAIL mode who did not receive a phone call, information about the household composition was extracted from the questionnaires by members of the SOEP team at Kantar. In CAWI mode, the team that called households from a telephone studio entered contact data, reasons for a refusal to participate and household information into a database screen.

## **Showcards**

Showcards were used by CAPI interviewers during the interview to show respondents longer item lists or scales.

## **Project instruction book**

In addition to a shorter project description, interviewers in CAPI and PAPI mode received a highly detailed instruction manual that was about 75 pages long. This manual contained information on special features of the current wave, specific processing instructions and questionnaires as well as background information on the project. There were two different versions of this manual – one for samples A-Q without L2/3 and another one for L2/3. The phone interviewers in MAIL and CAWI mode who called to update household information received a short instruction manual (MAIL) or a handout of the slides shown during interviewer training (CAWI).

## **Other interviewer material**

In addition, the interviewers got contact cards for households that could not be reached at home and household information cards with information on individual numbers, names, years of birth, types of questionnaires, incentives, survey modes and notes.

## 1.5 Conducting the survey

### 1.5.1 Survey mode

Since the SOEP was launched in 1984, face-to-face interviewing has been the primary method of data collection. Up to the year 2000, all face-to-face interviews were conducted using paper-and-pencil interviews (**PAPI**). Since then, SOEP interviewers have gradually started using laptops to conduct interviews in **CAPI** (computer-assisted personal interviewing). Since sample J in 2011, CAPI has been used exclusively to interview the respondents from refresher samples. However, PAPI is still used to collect data from the respondents in samples A–H if they prefer, or if their interviewer does not have a laptop.

Another method of interviewing is used in multi-person households from samples A–H. Individuals who were unable to provide an interview while the interviewer was present are offered the option of completing a paper questionnaire on their own as a means of reducing partial unit non-response (PUNR). The option of interviewing more than one person simultaneously with the help of paper questionnaires can be useful for reducing the overall length of interviewer visits to households with many members, thereby increasing participation. This method is a mixture of face-to-face interviewing and self-administered interviewing (**SELF**). Although this option is actually an exception, the longer a sample exists, the more frequently it is used to ensure low PUNR in larger households.

Another type of fieldwork processing used exclusively in core samples A–H is known as “central administration of fieldwork” (**MAIL**) in which respondents complete their questionnaires at home and return them by mail. This was first used as a refusal conversion process in the second wave of the SOEP in 1985 and is focused on households that did not agree to any further visits from an interviewer or could not be convinced by interviewers to participate for other reasons. As part of this process, households are contacted by telephone and asked to keep participating in the study. If this is successful, basic household information is collected or updated, and the questionnaires are sent by mail. Thus, in these households, questionnaires are fully self-administered. This mode shift often leads to a conversion of soft refusals, in turn improving the stability of long-term samples A–H.

The last addition to the SOEP in terms of survey modes is computer-assisted web interviewing (**CAWI**); this was introduced when samples L2/3 were integrated into the SOEP in 2014. The samples switched from an exclusively interviewer-assisted mode to a CATI/CAWI hybrid approach, followed by CAPI. The aim in every wave is, on the one hand, to recruit as many households as possible for participation by internet, and on the other hand, to maintain a high panel stability rate. The gross sample is thus divided into various subgroups depending on the mode of participation in previous years. Households that participated online at least once since 2014 were processed online first in 2020. This includes households that participated in CAPI in 2017 but did not explicitly refuse to do the interviews online. A CAPI interviewer was immediately sent to households that rejected the CAWI mode in previous waves. Households that did not answer the CAWI questionnaires during the first three months of CAWI fieldwork were sent a CAPI interviewer as well. In order to reduce both potential qualitative disadvantages and negative response-rate effects of using CAWI instead of CAPI, CATI interviewers contacted each household in the CAWI population to encourage online participation. They also made a list of all household members to ensure that the right set of CAWI questionnaires would be provided.

However, the year 2020 was the year of the coronavirus pandemic, and this posed a challenge for a survey like the SOEP, which is conducted primarily through face-to-face interviewing. The solution was to lift the usually rigorous mode restrictions described above to allow for more flexibility. This resulted in an unusually high number of SELF and PAPI interviews in the newer samples J–Q as well as the addition of CAPI by telephone (CAPI-TEL) to the mix of interview modes used in the SOEP.

The following tables display the interviewing modes applied for each respective questionnaire while separating different groups of samples. **Table 1.9** shows that in 2020 39.1 percent of all individual interviews were conducted via CAPI (standard with another 13.6 percent of CAPI via telephone, compared to 68.2 percent in 2019). Apart from that, paper-and-pencil interviews (PAPI) are still conducted with a small number of people, especially from the older samples A-H (2.4 percent). Some participants prefer to answer the questionnaires on their own – either with or without the interviewer present (SELF, 31.0 percent) this was especially true in 2020. 9.5 percent of all individual questionnaires in 2020 were sent to the panel members via mail, completed and then returned (MAIL). Finally, 5.5 percent of all individual interviews in 2020 were conducted via web interviewing (CAWI). This mode is used exclusively for respondents in sample L2/3. Within this sample, 40.9 percent of all individual questionnaires in 2020 were completed in CAWI mode.

**Table 1.9: Interviewing modes by sub-samples – individual questionnaire**

Individual level	Interviewer-Based				Centrally Administered		Total	
	CAPI-TEL	CAPI	PAPI	SELF	MAIL	CAWI		
A-H	Abs.	453	1,709	246	2,931	2,026	0	7,365
	In %	6.2	23.2	3.3	39.8	27.5	0.0	100.0
J, K, L1, N, O	Abs.	1,433	4,737	20	2,629	0	0	8,819
	In %	16.2	53.7	0.2	29.8	0.0	0.0	100.0
L2/3	Abs.	185	1,420	0	80	0	1,168	2,853
	In %	6.5	49.8	0.0	2.8	0.0	40.9	100.0
P	Abs.	569	327	21	787	0	0	1,704
	In %	33.4	19.2	1.2	46.2	0.0	0.0	100.0
Q	Abs.	251	124	1	164	0	0	540
	In %	46.5	23.0	0.2	30.4	0.0	0.0	100.0
Total	Abs.	<b>2,891</b>	<b>8,317</b>	<b>288</b>	<b>6,591</b>	<b>2,026</b>	<b>1,168</b>	<b>21,281</b>
	In %	<b>13.6</b>	<b>39.1</b>	<b>1.4</b>	<b>31.0</b>	<b>9.5</b>	<b>5.5</b>	<b>100.0</b>

The older samples A-H are characterized by considerable variation regarding the different modes (**Table 1.9**). CAPI as the primary interviewing mode is only used 23.2 percent of all individual interviews in this group of samples, CAPI via telephone (CAPI-TEL) was used in 6.2 percent of cases. SELF and MAIL interviews are almost as common in these samples. By contrast, for the newer samples J, K, L1, N, and O CAPI is usually the default mode used. However, in 2020 only 53.7 percent of individual interviews were conducted in standard CAPI, another 16.2 percent were conducted via telephone and SELF-interviewing accounted for another 29.8 percent of interviews. In samples P and Q, the shift to CAPI-TEL and SELF-interviewing was even more pronounced.



Table 1.10: Interviewing modes by sub-samples – household questionnaire

Household level	Interviewer-Based				Centrally Administered			Total
	CAPI-TEL	CAPI	PAPI	SELF	MAIL	CAWI		
A-H	Abs.	299	1,190	178	1,678	1,316	0	4,661
	In %	6.4	25.5	3.8	36.0	28.2	0.0	100.0
J, K, L1, N, O	Abs.	952	3,091	10	1,446	0	0	5,499
	In %	17.3	56.2	0.2	26.3	0.0	0.0	100.0
L2/3	Abs.	110	785	0	28	0	613	1,536
	In %	7.2	51.1	0.0	1.8	0.0	39.9	100.0
P	Abs.	446	272	16	477	0	0	1,211
	In %	36.8	22.5	1.3	39.4	0.0	0.0	100.0
Q	Abs.	201	105	1	113	0	0	420
	In %	47.9	25.0	0.2	26.9	0.0	0.0	100.0
Total	Abs.	<b>2,008</b>	<b>5,443</b>	<b>205</b>	<b>3,742</b>	<b>1,316</b>	<b>613</b>	<b>13,327</b>
	In %	<b>15.1</b>	<b>40.8</b>	<b>1.5</b>	<b>28.1</b>	<b>9.9</b>	<b>4.6</b>	<b>100.0</b>

Regarding the household questionnaires (**Table 1.10**), the distribution over the different modes is comparable to the one presented for the individual questionnaires. For the youth questionnaires (**Table 1.11**) and the child questionnaires (**Table 1.12**), the share of the different modes is also relatively similar, at least in the total sample.

Table 1.11: Interviewing modes by sub-samples – youth questionnaires

Individual level	Interviewer-Based				Centrally Administered			Total
	CAPI-TEL	CAPI	PAPI	SELF	MAIL	CAWI		
A-H	Abs.	1	40	6	83	62	0	192
	In %	0.5	20.8	3.1	43.2	32.3	0.0	100.0
J, K, L1, N, O	Abs.	15	291	8	293	2	0	609
	In %	2.5	47.8	1.3	48.1	0.3	0.0	100.0
L2/3	Abs.	1	191	0	27	0	159	378
	In %	0.3	50.5	0.0	7.1	0.0	42.1	100.0
P	Abs.	1	12	1	51	0	0	65
	In %	1.5	18.5	1.5	78.5	0.0	0.0	100.0
Q	Abs.	1	2	0	7	0	0	10
	In %	10.0	20.0	0.0	70.0	0.0	0.0	100.0
Total	Abs.	<b>19</b>	<b>536</b>	<b>15</b>	<b>461</b>	<b>64</b>	<b>159</b>	<b>1.254</b>
	In %	<b>1.5</b>	<b>42.7</b>	<b>1.2</b>	<b>36.8</b>	<b>5.1</b>	<b>12.7</b>	<b>100.0</b>

Table 1.12: Interviewing modes by sub-samples – child questionnaires

Individual level	Interviewer-Based				Centrally Administered			Total
	CAPI-TEL	CAPI	PAPI	SELF	MAIL	CAWI		
A-H	Abs.	1	58	7	55	128	0	249
	In %	0.4	23.3	2.8	22.1	51.4	0.0	100.0
J, K, L1, N, O	Abs.	18	437	7	218	3	0	683
	In %	2.6	64.0	1.0	31.9	0.4	0.0	100.0
L2/3	Abs.	0	73	0	4	0	43	120
	In %	0.0	60.8	0.0	3.3	0.0	35.8	100.0
P	Abs.	8	38	0	37	1	0	84
	In %	9.5	45.2	0.0	44.0	1.2	0.0	100.0
Q	Abs.	0	5	0	6	0	0	11
	In %	0.0	45.5	0.0	54.5	0.0	0.0	100.0
Total	Abs.	27	611	14	320	132	43	1.147
	In %	2.4	53.3	1.2	27.9	11.5	3.7	100.0

## 1.5.2 Fieldwork timing

For samples A-H, J-K, L1, N, O and P data collection covered a period of roughly seven months. Fieldwork started at the end of January for samples A-O and one month later for sample P, while the last interviews were conducted in early September. Fieldwork for sample Q started in March and ended in August. Within the first three full months, 66.1 percent of all net interviews for samples A-O were completed, as indicated by the figures in **Table 1.13**. For sample L2/3, data collection started in July and was finalized in December. Within the first three months, 65.9 percent of all interviews were completed in this sample. For all samples, the last months of fieldwork are dedicated to contacting difficult-to-reach households and respondents, which includes:

- Households requiring tracing of new addresses
- New households
- Respondents who are difficult to reach due to work-related absence, a stay in a foreign country or an ongoing illness
- Respondents who refuse their participation due to lack of time

The extensive follow-up processing during the final months of the fieldwork period is substantial in the effort to minimize panel mortality and maintain high levels of panel stability in all samples.

Table 1.13: Monthly fieldwork progress by sub-samples

Household Level	A-H, J-K, L1, N, O				L2/3				P				Q			
	Gross Sample		Net Sample		Gross Sample		Net Sample		Gross Sample		Net Sample		Gross Sample		Net Sample	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
January <sup>1</sup>	172	1.4	3	0.0	-	-	-	-	-	-	-	-	-	-	-	-
February	3,947	32.6	3,602	35.1	-	-	-	-	53	2.7	12	1.0	-	-	-	-
March	3,453	28.5	3,184	31.0	-	-	-	-	283	14.3	194	15.8	190	39.4	178	42.1
April	2,241	18.5	1,972	19.2	-	-	-	-	309	15.6	235	19.1	140	29.0	130	30.7
May	1,333	11.0	1,010	9.8	-	-	-	-	501	25.3	344	28.0	92	19.1	81	19.1
June	497	4.1	298	2.9	-	-	-	-	342	17.2	235	19.1	29	6.0	22	5.2
July	339	2.8	156	1.5	244	12.3	228	14.8	281	14.2	145	11.8	25	5.2	9	2.1
August	122	1.0	43	0.4	402	20.3	362	23.5	157	7.9	51	4.1	6	1.2	3	0.7
September	5	0.0	2	0.0	489	24.7	424	27.6	57	2.9	13	1.1	-	-	-	-
October	-	-	-	-	430	21.8	333	21.7	-	-	-	-	-	-	-	-
November	-	-	-	-	358	18.1	179	11.6	-	-	-	-	-	-	-	-
December	-	-	-	-	53	2.7	12	0.8	-	-	-	-	-	-	-	-

<sup>1</sup> Including households that refused to take part in the survey prior to the start of fieldwork.

## 1.5.3 Translations

Interviewers receive translated paper questionnaires that can be used as an aid when conducting the interview via CAPI or PAPI mode. The questionnaire for households as well as the individual and life-history questionnaires are available in five foreign languages: English, Russian, Turkish, Polish and Romanian (Table 1.14).

Table 1.14: Translated paper questionnaires

	English	Russian	Turkish	Polish	Romanian
Household questionnaire	✓	✓	✓	✓	✓
Individual questionnaire	✓	✓	✓	✓	✓
Life-history questionnaire	✓	✓	✓	✓	✓
Youth questionnaires	-	-	-	-	-
Mother and child questionnaires/ questionnaire for parents	-	-	-	-	-
Additional questionnaires <sup>1</sup>	-	-	-	-	-

<sup>1</sup> Gap, Deceased person.

## 1.5.4 Panel maintenance and incentives

Before fieldwork began, all households from samples A-Q received a letter announcing the beginning of the new wave. For almost all households in samples A-H, the letter included a lottery ticket as an unconditional incentive for every adult who participated in the last wave. Due to an experiment that was conducted a few years ago, a group of approx. 600 households in samples A-H received a cash incentive instead (**Table 1.15**). Participants in the newer samples J, K, L1, N, O, and Q received cash incentives as well. The individual questionnaire was rewarded with 10 euros, while the shorter household questionnaire was rewarded with 5 euros. Respondents in sample P received a lottery ticket.

Participants in sample L2/3 earned 5 euros for every questionnaire, including the youth and child questionnaire. Additionally, a household received a bonus payment of 10 euros if all individual questionnaires have been completed. Participants who completed the survey in CAWI mode received their incentive as vouchers by mail or e-mail depending on personal preference. For CAPI participants, the interviewer paid the incentive in cash.

Table 1.15: Incentives samples A-Q

	A-H	J, K, L1, N, O, Q, P	L2/3
Incentives for adults	Lottery ticket: 4,927 hh Cash <sup>1</sup> : 602 hh	HH: 5 euros Adult: 10 euros Sample P: Lottery ticket	HH: 5 euros Adult: 5 euros Bonus payment: 10 euros
Incentives for youth and child questionnaire	Youth quest.: age 16 or 17: Power bank Youth quest.: age 13 or 14: USB-stick Youth quest.: age 11 or 12: CAPI/PAPI: Bicycle repair kit; MAIL: Puzzle Child quest.: No incentive		All youth quest.: 5 euros All child quest.: 5 euros

<sup>1</sup> As part of an incentive experiment: same cash incentives as in samples J-Q; participants in MAIL mode received a check.

In addition to the individual incentives, interviewers brought a small gift to all households which were presented upon arrival. This year's gift was a high-quality branded shopping bag. Furthermore, parents of newborns received a small photo album via mail. Neither of these gifts were given to households in samples L2/3.

## 1.5.5 Movers and tracing

**Table 1.16** provides figures for the number of households requiring address inquiry including movers as well as new households. For this year's survey wave, 1,514 households (9.1 percent) were subject to address inquiry. 1,259 of these households were successfully traced, which amounts to a success rate of 83.2 percent. The most important source of information about addresses were interviewers (51.0 percent), followed by the postal service (17.7 percent) that provide information about residents' current addresses if mail is undeliverable and movers notify them of their new address. Sometimes, participants themselves let us know about a move. This source amounts to only 10.2 percent of address information about movers and new households.

**Table 1.16: Movers and sources of new addresses of administered sample 2020**

Household Level	Total		A-H, J-K, L1, N, O		L2/3		P		Q	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample</b>	16,550	100.0	11,312	100.0	1,976	100.0	797	100.0	1,983	100.0
Movers and new households	1,514	9.1	970	8.6	341	17.3	79	9.9	79	4.0
Success tracing										
Tracing successful	1,259	83.2	846	87.2	237	69.5	66	83.5	68	86.1
Tracing not successful	255	16.8	124	12.8	104	30.5	13	16.5	11	13.9
Source										
Interviewer	773	51.0	524	54.0	149	43.7	45	57.0	30	38.0
Postal service	268	17.7	170	17.5	59	17.3	8	10.1	19	24.1
Local registration offices	155	10.2	126	13.0	0	0.0	16	20.3	7	8.9
Participant	155	10.2	74	7.6	75	22.0	1	1.3	3	3.8
Post-Address Research	164	10.8	77	7.9	58	17.0	9	11.4	20	25.3

## 1.5.6 Interviewer characteristics, training & monitoring

For samples A-Q, a total of 485 interviewers were active in 2020 (**Table 1.17**). Among those interviewers, gender is distributed relatively evenly with a slightly higher presence of male interviewers. Interviewer age is relatively high, with 70.3 percent of all interviewers being between 60 and 79 years old. This not only holds true for the older samples A-H but also for the more recently established samples. Most interviewers processed between 20 and 99 households this year, but there are also 87 interviewers who were responsible for fewer than five households.

# KANTAR PUBLIC

Table 1.17: Interviewer characteristics

Interviewer Level	Total		A-H		J-K		L1		L2/3		N		O		P		Q	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Number of Interviewers</b>	485	100.0	379	100.0	265	100.0	175	100.0	225	100.0	228	100.0	124	100.0	169	100.0	175	100.0
<b>Gender</b>																		
Male	278	57.3	206	54.4	149	56.2	103	58.9	124	55.1	126	55.3	65	52.4	98	58.0	98	56.0
Female	207	42.7	173	45.6	116	43.8	72	41.1	101	44.9	102	44.7	59	47.6	71	42.0	77	44.0
<b>Age</b>																		
21-39 years	9	1.9	4	1.1	2	0.8	2	1.1	2	0.9	2	0.9	2	1.6	3	1.8	5	2.9
40-59 years	91	18.8	65	17.2	42	15.8	35	20.0	33	14.7	50	21.9	30	24.2	41	24.3	37	21.1
60-79 years	341	70.3	267	70.4	194	73.2	121	69.1	168	74.7	159	69.7	83	66.9	117	69.2	120	68.6
80+ years	44	9.1	43	11.3	27	10.2	17	9.7	22	9.8	17	7.5	9	7.3	8	4.7	13	7.4
<b>Number of households (gross)</b>																		
Fewer than 5	87	17.9	153	40.4	90	34.0	99	56.6	120	53.3	83	36.4	62	50.0	52	30.8	155	88.6
5 – 19	169	34.8	175	46.2	142	53.6	72	41.1	100	44.4	118	51.8	58	46.8	99	58.6	17	9.7
20 – 99	213	43.9	51	13.5	33	12.5	4	2.3	5	2.2	27	11.8	4	3.2	18	10.7	3	1.7
More than 99	16	3.3	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0

**Table 1.18** gives an overview of interviewer experience, both with Kantar overall and with the SOEP specifically. Interviewers with ten and more years of experience make up the majority across all sub-samples and are particularly present in samples A-H. 30 Interviewers have been conducting interviews for the SOEP (almost) since the beginning of the study.



# KANTAR PUBLIC

Table 1.18: Interviewer experience

Interviewer Level	Total		A-H		J-K		L1		L2/3		N		O		P		Q	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Number of Interviewers</b>	485	100.0	379	100.0	265	100.0	175	100	225	100	228	100.0	124	100.0	175	100.0	169	100.0
<b>Experience with Kantar</b>																		
0-4 years	84	17.3	36	9.5	29	10.9	22	12.6	27	12.0	33	14.5	34	27.4	34	19.4	43	25.4
5-9 years	81	16.7	59	15.6	50	18.9	27	15.4	39	17.3	45	19.7	23	18.5	30	17.1	39	23.1
10-19 years	159	32.8	133	35.1	112	42.3	89	50.9	103	45.8	94	41.2	46	37.1	74	42.3	65	38.5
20-29 years	84	17.3	77	20.3	42	15.8	21	12.0	32	14.2	29	12.7	13	10.5	21	12.0	12	7.1
30-39 years	47	9.7	45	11.9	21	7.9	9	5.1	16	7.1	18	7.9	5	4.0	11	6.3	5	3.0
More than 40 years	30	6.2	29	7.7	11	4.2	7	4.0	8	3.6	9	3.9	3	2.4	5	2.9	5	3.0
<b>Experience with SOEP<sup>1</sup></b>																		
0-4 years	104	22.4	50	13.9	42	16.0	33	18.9	36	16.1	46	20.3	38	30.6	44	25.1	52	30.8
5-9 years	137	29.5	102	28.4	95	36.1	48	27.4	79	35.3	75	33.0	36	29.0	59	33.7	56	33.1
10-19 years	120	25.8	108	30.1	81	30.8	67	38.3	72	32.1	66	29.1	34	27.4	50	28.6	45	26.6
20-29 years	68	14.6	64	17.8	30	11.4	18	10.3	26	11.6	27	11.9	11	8.9	15	8.6	11	6.5
30-34 years	6	1.3	6	1.7	4	1.5	3	1.7	4	1.8	4	1.8	2	1.6	3	1.7	2	1.2
More than 35 years	30	6.5	29	8.1	11	4.2	6	3.4	7	3.1	9	4.0	3	2.4	4	2.3	3	1.8

<sup>1</sup>For 22 interviewers the experience with SOEP surveys is unknown.

For the interviewer training in January 2020, the so-called “contact interviewer training format” was used again. In a central, one-day event, members of the SOEP team at Kantar, the field organization and representatives from the DIW trained contact interviewers who then in turn trained the interviewers in the regions for which they are responsible.

Contents of the training were:

- Scope, timing and procedure of the various SOEP studies in 2020
- Importance of high response rates and low partial unit non-response for the SOEP surveys
- Special features of the 2019 survey year:
  - Record linkage with the German statutory pension insurance
  - Online survey “Self-Employment and Entrepreneurship in Germany”
  - Item-Non-Response
  - Second wave of sample P

Kantar places high priority on interviewer monitoring and has put an ISO-certificated process in place that is audited regularly. Kantar adheres to the German Business Association of Market and Social Research Institutes (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V., ADM) standards for internal regulation and monitoring of all systems and procedures. This means that a minimum of 10 percent of Kantar’s annual interviews are checked, and every interviewer is monitored at least once a year.

In 2018, we expanded these existing quality-control measures for interviewer monitoring in the SOEP projects. The basic interviewer monitoring concept was expanded to a project-based control system for all SOEP samples: A number of participating households is contacted shortly after the interview by letter or phone asking them to confirm their participation in a regularly conducted interview. In case of inconsistencies and/or irregularities, we attempt to gain clarification through direct contact with respondents, primarily by telephone. We have also begun building a comprehensive suite of bespoke analytical tools to uncover irregularities in the survey data and para data. The indicators produced by this tool are can led to further measures taken in the re-contacting process.

In 2019, we have intensified our interviewer monitoring processes by also using para data and interview data in all SOEP samples starting soon after fieldwork begins. Results from these data checks have been shared regularly with the SOEP team at DIW Berlin to discuss what steps to take with interviewers who show irregularities and to develop indicators. In 2019 we have started offering respondents an option to answer the short questionnaire online as part of the re-contacting process instead of the paper questionnaire.

**Table 1.19** shows details for the re-contacting process in samples A-Q. In existing panel samples, the aim is to re-contact at least 10 percent of households for every interviewer in every SOEP project. However, the average share of re-contacted households in A-Q was significantly higher, at 28.1 percent in 2020. There are several reasons for this. First, for interviewers who are responsible for only a few households, more than 10 percent of their households were re-contacted to generate a sufficient response. Second, for interviewers with low response rates from re-contacted households, the number of re-contacted households was sometimes increased. In terms of survey mode, a short paper questionnaire was followed up by phone for interviewers with a response rate significantly below 30 percent. Third, we decided to recontact even more households due to the unusual mode-mix that was adopted because of the coronavirus pandemic (see section 1.5.1)

Overall, we re-contacted 3,353 households in samples A-Q and received feedback from 1,421 households. This results in a response rate of 42.4 percent overall. Using the results from the re-contacting process, we identified one interviewer who had not adhered to our standards in conducting interviews in the 2020 fieldwork

period. The data was flagged, and the identifier was communicated to the DIW-SOEP for further examination and processing.

Table 1.19: **Interviewer monitoring**

	Abs.	In % (net sample)	In % (re-contacted household)
<b>By households</b>			
Households in net sample <sup>1</sup>	11,922	100.0	-
Re-contacted households	3,353	28.1	100.0
Households with feedback	1,421	11.9	42.4
<b>By interviewers</b>			
Interviewers in net sample <sup>1</sup>	461	100.0	-
Interviewers with re-contacted households	448	97.2	-
Interviewers with non-standard behavior	1	0.2	-

<sup>1</sup> Interviewer based, A-Q without L2/3.

## 1.6 Fieldwork results

### 1.6.1 Participation on household level

**Table 1.20** displays participation and reasons for dropping out for three different types of households. Of 16,550 households in samples A-Q, 13,460 households took part in the survey in 2020 (81.3 percent).

Table 1.20: **Participation by type of household (with AAPOR codes)**

	Total		Respondents in previous wave		Dropouts in previous wave		New households	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
Gross Sample	16,550	100.0	14,924	100.0	1,132	100.0	494	100.0
<b>Interview (1.0)</b>	<b>13,460</b>	<b>81.3</b>	<b>12,770</b>	<b>85.6</b>	<b>460</b>	<b>40.6</b>	<b>230</b>	<b>46.6</b>
Complete (1.1)	10,645	64.3	10,104	67.7	343	30.3	198	40.1
Partial (1.2)	2,815	17.0	2,666	17.9	117	10.3	32	6.5
<b>Non-Interview (2.0)</b>	<b>3,090</b>	<b>18.7</b>	<b>2,154</b>	<b>14.4</b>	<b>672</b>	<b>59.4</b>	<b>264</b>	<b>53.4</b>
<b>Temporary dropout</b>	<b>1,603</b>	<b>9.7</b>	<b>1,218</b>	<b>8.2</b>	<b>267</b>	<b>23.6</b>	<b>118</b>	<b>23.9</b>
Non-contact (2.20)	559	3.4	356	2.4	162	14.3	41	8.3
Temporary refusal (2.351; 2.353; 2.354)	885	5.3	760	5.1	74	6.5	51	10.3
Temporarily physically or mentally unable/incompetent (2.321)	78	0.5	71	0.5	5	0.4	2	0.4
Household could not be traced (temporary) (3.18; 2.4)	40	0.2	12	0.1	14	1.2	14	2.8
Other temp. (2.52)	39	0.2	19	0.1	10	0.9	10	2.0
Household level language problems (2.331)	2	0.0	0	0.0	2	0.2	0	0.0
<b>Final Dropout</b>	<b>1,487</b>	<b>9.0</b>	<b>936</b>	<b>6.3</b>	<b>405</b>	<b>35.8</b>	<b>146</b>	<b>29.6</b>
Permanent refusal (2.111)	1,220	7.4	768	5.1	355	31.4	97	19.6
Permanently physically or mentally unable/incompetent (2.322)	88	0.5	66	0.4	19	1.7	3	0.6
Deceased (2.31)	61	0.4	53	0.4	8	0.7	0	0.0
Moved abroad (4.2)	19	0.1	14	0.1	5	0.4	0	0.0
Household dissolved (4.2)	23	0.1	18	0.1	5	0.4	0	0.0
Household untraceable (4.4)	76	0.5	17	0.1	13	1.1	46	9.3
Dropped out temporarily in two consecutive waves	0	0.0	0	0.0	0	0.0	0	0.0

Of all households that participated in the previous year, 85.6 percent were successfully recruited again in this year's wave. Additionally, 460 households which had dropped out during the previous year were reintegrated into the survey (40.6 percent). Of all new households, 46.6 percent were successfully converted into participating households. 3,090 households (18.7 percent of the gross sample) did not take part in the current wave. 51.9 percent of these households dropped out temporarily and will be contacted again next year. The remaining dropouts are permanent, so these households will not be contacted again. The main reason for final dropouts by far is permanent refusal, while other reasons do not account for many dropouts.

# KANTAR PUBLIC

Table 1.21: Participation by sample (with AAPOR codes)

	Total		A-H		J-K		L1		L2/3		N		O		P		Q	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	Abs.	In %	In %	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
Gross Sample	16,550	100.0	5,529	100.0	2,599	100.0	1,029	100.0	1,976	100.0	2,155	100.0	797	100.0	1,983	100.0	482	100.0
Interview (1.0)	13,460	81.3	4,726	85.5	2,265	87.1	866	84.2	1,538	77.8	1,844	85.6	569	71.4	1,229	62.0	423	87.8
Complete (1.1)	10,645	64.3	3,963	71.7	1,936	74.5	770	74.8	1,220	61.7	1,359	63.1	489	61.4	588	29.7	320	66.4
Partial (1.2)	2,815	17.0	763	13.8	329	12.7	96	9.3	318	16.1	485	22.5	80	10.0	641	32.3	103	21.4
Non-Interview (2.0)	3,090	18.7	803	14.5	334	12.9	163	15.8	438	22.2	311	14.4	228	28.6	754	38.0	59	12.2
Temporary dropout	1,603	9.7	268	4.8	171	6.6	97	9.4	272	13.8	164	7.6	152	19.1	437	22.0	42	8.7
Non-contact (2.20)	559	3.4	49	0.9	62	2.4	35	3.4	119	6.0	68	3.2	82	10.3	127	6.4	17	3.5
Temporary refusal (2.351; 2.353; 2.354)	885	5.3	190	3.4	86	3.3	57	5.5	103	5.2	80	3.7	55	6.9	295	14.9	19	3.9
Temporarily physically or mentally unable/incompetent (2.321)	78	0.5	19	0.3	17	0.7	2	0.2	6	0.3	12	0.6	6	0.8	11	0.6	5	1.0
Household could not be traced (temporary) (3.18; 2.4)	40	0.2	4	0.1	5	0.2	1	0.1	17	0.9	4	0.2	6	0.8	2	0.1	1	0.2
Other temp. (2.52)	39	0.2	6	0.1	1	0.0	2	0.2	27	1.4	0	0.0	1	0.1	2	0.1	0	0.0
Household level language problems (2.331)	2	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	2	0.3	0	0.0	0	0.0
<b>Final Dropout</b>	<b>1,487</b>	<b>9.0</b>	<b>535</b>	<b>9.7</b>	<b>163</b>	<b>6.3</b>	<b>66</b>	<b>6.4</b>	<b>166</b>	<b>8.4</b>	<b>147</b>	<b>6.8</b>	<b>76</b>	<b>9.5</b>	<b>317</b>	<b>16.0</b>	<b>17</b>	<b>3.5</b>
Permanent refusal (2.111)	1,220	7.4	447	8.1	103	4.0	62	6.0	108	5.5	132	6.1	60	7.5	295	14.9	13	2.7
Permanently physically or mentally unable (2.322)	88	0.5	39	0.7	27	1.0	1	0.1	2	0.1	3	0.1	7	0.9	8	0.4	1	0.2
Deceased (2.31)	61	0.4	33	0.6	16	0.6	0	0.0	2	0.1	2	0.1	5	0.6	1	0.1	2	0.4
Moved abroad (4.2)	19	0.1	4	0.1	1	0.0	1	0.1	3	0.2	3	0.1	3	0.4	4	0.2	0	0.0
Household dissolved (4.3)	23	0.1	4	0.1	6	0.2	1	0.1	8	0.4	3	0.1	1	0.1	0	0.0	0	0.0
Household untraceable (4.4)	76	0.5	8	0.1	10	0.4	1	0.1	43	2.2	4	0.2	0	0.0	9	0.5	1	0.2
Dropped out temporarily in two consecutive waves	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0

**Table 1.21** displays fieldwork results on the household level for different samples and sample groups. The participation rate was slightly higher for the samples A-H, J-K, L1, N, and Q than for CAWI/CAPI samples L2/3, O, and for the newest sample P.

**Table 1.22: Participation by type of fieldwork (with AAPOR codes)<sup>1</sup>**

	Total		Interviewer-based		MAIL <sup>2</sup>		CAWI <sup>3</sup>	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
Gross Sample	16,550	100.0	14,270	100.0	1,535	100.0	745	100.0
<b>Interview (1.0)</b>	<b>13,460</b>	<b>81.3</b>	<b>11,596</b>	<b>81.3</b>	<b>1,277</b>	<b>83.2</b>	<b>587</b>	<b>78.8</b>
Complete (1.1)	10,645	64.3	9,205	64.5	1,008	65.7	432	58.0
Partial (1.2)	2,815	17.0	2,391	16.8	269	17.5	155	20.8
<b>Non-Interview (2.0)</b>	<b>3,090</b>	<b>18.7</b>	<b>2,674</b>	<b>18.7</b>	<b>258</b>	<b>16.8</b>	<b>158</b>	<b>21.2</b>
<b>Temporary dropout</b>	<b>1,603</b>	<b>9.7</b>	<b>1,391</b>	<b>9.7</b>	<b>99</b>	<b>6.4</b>	<b>113</b>	<b>15.2</b>
Non-contact (2.20)	559	3.4	513	3.6	2	0.1	44	5.9
Temporary refusal (2.351; 2.353; 2.354)	885	5.3	734	5.1	95	6.2	56	7.5
Temporarily physically or mentally unable/incompetent (2.321)	78	0.5	76	0.5	1	0.1	1	0.1
Household could not be traced (temporary) (3.18; 2.4)	40	0.2	33	0.2	1	0.1	6	0.8
Other temp. (2.52)	39	0.2	33	0.2	0	0.0	6	0.8
Household level language problems (2.331)	2	0.0	2	0.0	0	0.0	0	0.0
<b>Final Dropout</b>	<b>1,487</b>	<b>9.0</b>	<b>1,283</b>	<b>9.0</b>	<b>159</b>	<b>10.4</b>	<b>45</b>	<b>6.0</b>
Permanent refusal (2.111)	1,220	7.4	1,052	7.4	142	9.3	26	3.5
Permanently physically or mentally unable/incompetent (2.322)	88	0.5	82	0.6	6	0.4	0	0.0
Deceased (2.31)	61	0.4	58	0.4	3	0.2	0	0.0
Moved abroad (4.2)	19	0.1	15	0.1	2	0.1	2	0.3
Household dissolved (4.3)	23	0.1	18	0.1	2	0.1	3	0.4
Household untraceable (4.4)	76	0.5	58	0.4	4	0.3	14	1.9
Dropped out temporarily in two consecutive waves	0	0.0	0	0.0	0	0.0	0	0.0

<sup>1</sup> By mode in which the household participated in 2019; For new households, mode in which the old households participated in 2019.

<sup>2</sup> All households in the MAIL mode used to be dropouts in the interviewer-based mode and were willing to continue participation via mail.

<sup>3</sup> Households with start mode CAWI automatically switch to interviewer-based mode if they fail to participate online. Households from sample L2/3 with start mode interviewer can also switch to CAWI if they express interest in participating online.

**Table 1.22** presents the level of participation for the different types of fieldwork. Since the mode can be switched during the wave if the household does not take part in the initial mode, it is important to note that the

table relies on the final mode in which the household participated in the previous wave. So, a household that is interviewed in CAWI mode first in 2020 might have participated in the interviewer-based mode if it failed to fill out the online questionnaires. This was the case for 10.4 percent of the households that started in CAWI mode (**Table 1.23**).

With 83.2 percent, the highest level of participation is reached in the MAIL mode, followed by the interviewer-based modes that achieved 81.3 percent (**Table 1.22**). CAWI mode as used in sample L2/3 produced slightly lower levels of participation. In 2020, 78.8 percent of households took part in the study via CAWI mode. With 15.2 percent, the CAWI mode shows a higher tendency towards temporary dropouts compared to other types of fieldwork. The percentage of final dropouts in turn is lower for web interviews.

**Table 1.23: Type of fieldwork 2020 by type of fieldwork 2019<sup>1</sup>**

2020	2019		Interviewer-based		MAIL		CAWI	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Interview</b>	<b>13,460</b>	<b>100.0</b>	<b>11,596</b>	<b>100.0</b>	<b>1,277</b>	<b>100.0</b>	<b>587</b>	<b>100.0</b>
Interviewer-based	11,531	85.7	11,467	98.9	3	0.2	61	10.4
MAIL	1,316	9.8	42	0.4	1,274	99.8	0	0.0
CAWI	613	4.6	87	0.8	0	0.0	526	89.6

<sup>1</sup> By mode in which the household participated in 2018; For new households, mode in which the old households participated in 2018.

The overall response rate as well as the respective response rates for three different types of households are displayed in **Table 1.24**. To calculate the response rate, the total number of participating households is divided by the gross sample minus households which were impossible to survey in the present year. This includes households which moved abroad and households that permanently cannot be traced or because the last household member died. The overall response rate for this year's survey amounts to 82.2 percent. For households that participated in the previous wave, a response rate of 86.1 percent was reached. Response rates for the other two types of households, dropouts in the previous wave and new households, are generally lower. These amounted to 41.8 percent and 51.3 percent, respectively, in 2020.

**Table 1.24: Response rate by type of household (in percent)**

	Total	Respondents in previous wave	Dropouts in previous wave	New households
Response rate <sup>1</sup>	82.2	86.2	41.8	51.3

<sup>1</sup> RR = percentage of all households with at least one hh and individual interview in the gross sample (gross sample adjusted for households where the last person is deceased or the household moved abroad, is permanently untraceable or dissolved households where the last member moved into another SOEP household).

Response rates for different samples and sample groups are presented in **Table 1.25**. Sample groups A-H, J-K; N and Q all score response rates over 85 percent. With 84.4 percent for L1, 80.1 percent for L2/3, response rates are slightly lower for these samples. These results mirror the general tendency of older samples and interviewer-administered samples to gain higher response rates. Sample O and sample P are both on the path to consolidation still with response rate at 72.2 percent for O and 62.4 percent for sample P.



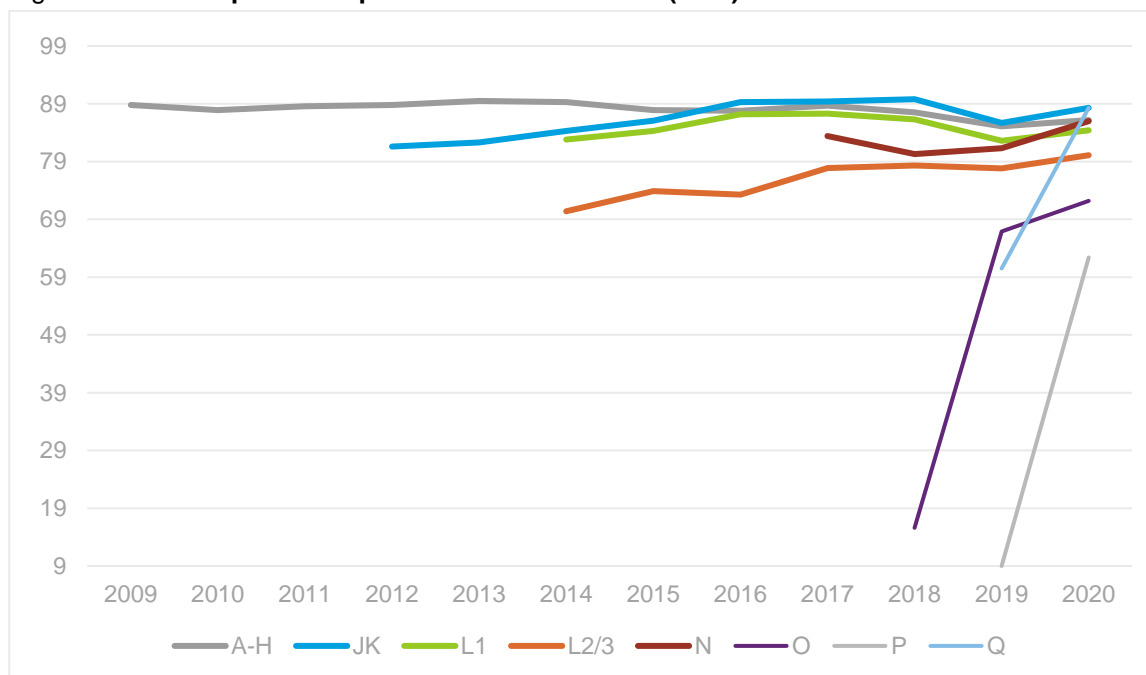
Table 1.25: **Response rate by sample (in percent)**

	Total	A-H	J-K	L1	L2/3	N	O	P	Q
Response rate <sup>1</sup>	82.2	86.2	88.3	84.4	80.1	86.0	72.2	62.4	88.3

<sup>1</sup> RR = percentage of all households with at least one hh and individual interview in the gross sample (gross sample adjusted for households where the last person is deceased or the household moved abroad, is permanently untraceable or dissolved households where the last member moved into another SOEP household).

As depicted in **Figure 1.3**, the response rates for samples A-H have been fairly stable between 2009 and 2020. After a continuous rise in response rates since their inclusion, samples J and K have been exceeding the response levels of the older samples since 2016. The response rates for L1 follow a similar path on a slightly lower level. Meanwhile, L2/3 started out on a much lower level than L1 in 2014 and increased by almost 10 percentage points in the following years. With 86.6 percent for sample N in 2020, the response rate increased again compared to the prior year (81.3 percent). Sample P and Q both showed significant upswings in the second wave, which is unsurprising due to the fact that households that fail to participate in the first wave are not included in the gross sample of the second wave.

Figure 1.3: **Development response rates<sup>1</sup> since 2009 (in %)**

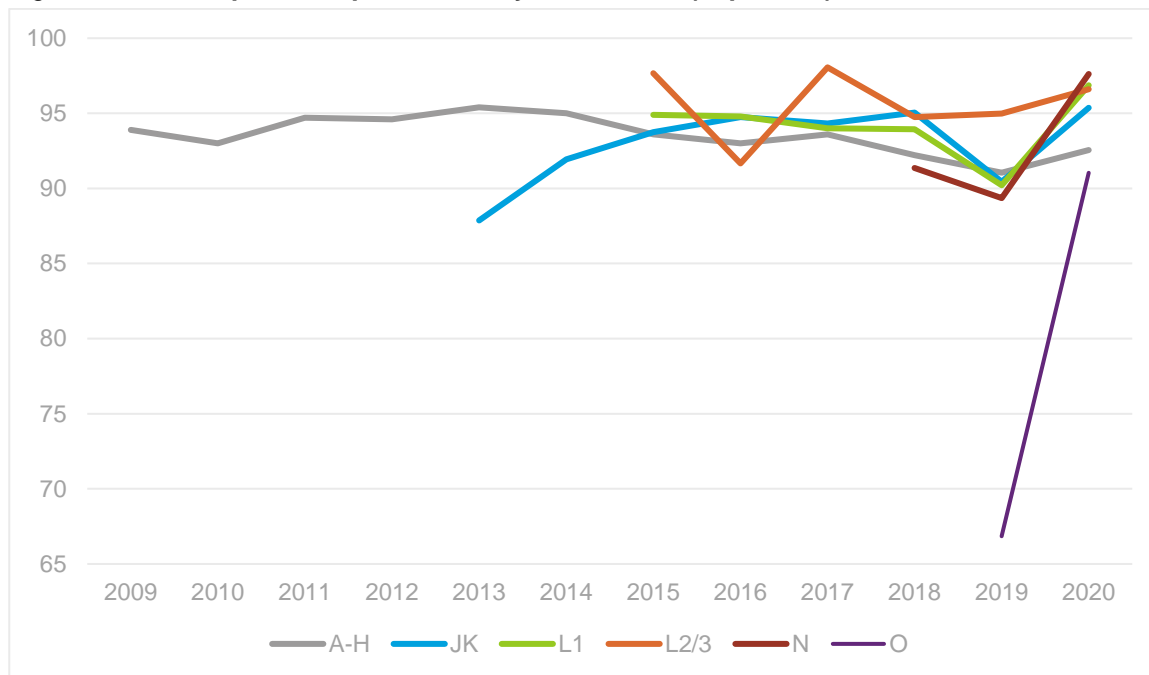


<sup>1</sup> RR= all households with at least one hh interview and one individual interview/(households in gross sample – households where the last person is deceased, moved abroad or is temporarily or permanently unable to participate or dissolved households where the last member moved into another SOEP household).

From a long-term perspective, panel stability can serve as an interesting indicator when monitoring and predicting a longitudinal sample's development in terms of overall size. Panel stability is calculated as the number of households participating in the current year compared to the number from the previous year. It reflects the net total effects of panel mortality on the one hand and panel growth (through new/split-off households or households that "re-joined" the panel) on the other hand.

In order to meaningfully assess panel stability rates over the years, the various subsamples should be processed for at least five consecutive waves. After this time period, the panel stability rates of samples are usually consolidated and therefore comparable. **Figure 1.4** shows that panel stability for sample groups A-H and J-K as well as for sample L1 has been relatively solid between 2009 and 2020. The level of stability slightly fluctuated between 2011 and 2013 for sample group A-H but never dropped below 90 percent. Panel stability for sample J-K shows a slight upward trend starting at 92 percent in 2014 and surpassing 95 percent in 2018 and 2020 after it had dropped to 90.4 percent in 2019. Panel stability for sample P was at 62.7 percent while sample Q reached a very good 88.7 percent panel stability.

Figure 1.4: Development of panel stability<sup>1</sup> since 2009 (in percent)



<sup>1</sup> Number of participating households divided by previous wave's net sample.

## 1.6.2 Participation on individual level

As presented in **Table 1.26**, 21,614 individuals from the core samples took part in the survey in 2020, which equals 63.8 percent of the gross sample. Consequently, 36.2 percent of all possible respondents in the gross sample did not participate this year. 19.3 percent of those individuals dropped out of the survey permanently, while 16.9 percent will be asked to participate again in next year's survey wave. Regarding all respondents from the previous wave, 84.9 percent took part in the present year. Of 5,869 temporary dropouts in 2019, 1,033 individuals were successfully reintegrated in 2019, 193 of those who had given a permanent refusal in an earlier wave and have been converted successfully. Additionally, 131 new panel members were recruited from existing or newly established households. Also, 567 young adults reached the age required to fully participate in the survey with the individual questionnaire. 59.3 percent of them did so.

Table 1.26: Participation by panel status (with AAPOR codes)

	Total		Respondents in previous wave		Temp. dropouts in previous wave		Permanent refusals in previous waves		New panel members <sup>1</sup>		Grown into Panel <sup>2</sup>	
	Abs.	In % gross sample	Abs.	In% gross sample	Abs.	In % gross sample	Abs.	In % gross sample	Abs.	In % gross sample	Abs.	In % gross sample
Gross Sample <sup>3</sup>	33,866	100.0	23,474	100.0	5,869	100.0	3,585	100.0	371	100.0	567	100.0
Interview (1.0)	21,614	63.8	19,921	84.9	1,033	17.6	193	5.4	131	35.3	336	59.3
Non-interview (2.0)	12,252	36.2	3,553	15.1	4,836	82.4	3,392	94.6	240	64.7	231	60.7
Temporary dropout	5,708	16.9	2,030	8.6	3,378	57.6	0	0.0	176	47.4	123	21.7
Non-contact (2.21)	83	0.2	43	0.2	20	0.3	0	0.0	10	2.7	10	1.8
Temporary refusal (2.112)	2,758	8.1	1,883	8.0	679	11.6	0	0.0	101	27.2	95	16.8
Temp. physically or mentally unable (2.321)	17	0.1	14	0.1	2	0.0	0	0.0	1	0.3	0	0.0
Person could not be traced (temporary) (3.18; 2.4)	46	0.1	21	0.1	23	0.4	0	0.0	0	0.0	2	0.4
Other temp. (2.52)	2,804	8.3	69	.3	2,654	45.2	0	0.0	64	17.3	16	2.8
Final Dropout	6,544	19.3	1,523	6.5	1,458	24.8	3,392	94.6	64	17.3	108	19.0
Permanent refusal (2.111)	6,234	18.4	1,316	5.6	1,363	23.2	3,392	94.6	62	16.7	102	18.0
Perm. physically or mentally unable (2.322)	22	0.1	11	0.0	8	0.1	0	0.0	2	0.5	1	0.2
Deceased (2.31)	150	0.4	130	0.6	19	0.3	0	0.0	0	0.0	1	0.2
Moved abroad (4.2)	33	0.1	22	0.1	11	0.2	0	0.0	0	0.0	0	0.0
Person untraceable (4.4)	105	0.3	44	0.2	57	1.0	0	0.0	0	0.0	4	0.7

<sup>1</sup> New household members who have never been part of the panel.

<sup>2</sup> Former youths who have been part of the panel and take part as official respondents for the first time (with the individual questionnaire).

<sup>3</sup> All household members intended to participate with the adult questionnaire in the current wave, not restricted to members of participating households.

The following **Table 1.27** includes the response rates for the different types of respondents. In total, an individual response rate of 64.4 percent was accomplished. For respondents in the previous wave, the response rate was 85.6 percent.

Table 1.27: **Response rate by panel status (in percent)**

	Total	Respondents in previous wave	Temp. dropouts in previous wave	Permanent refusals in previous waves	New panel members <sup>2</sup>	Grown into Panel <sup>3</sup>
Response rate <sup>1</sup>	64.4	85.6	17.9	5.4	35.3	59.8

<sup>1</sup> RR = percentage of all participants in the gross sample of individuals born before 2002 (gross sample adjusted for persons who are deceased, moved abroad or are untraceable).

<sup>2</sup> New household members who have never been part of the panel.

<sup>3</sup> Former youths who have been part of the panel and take part as official respondents for the first time (with the individual questionnaire).

Along with response rates, the rate of partial unit non-response (PUNR) serves as a critical indicator to measure the success of the fieldwork process in a household survey. Of 13,460 households that took part in the survey in 2020, 10,645 were not missing any interviews from adults (**Table 1.20**). In the SOEP, we measure PUNR only for households with more than one household member. These rates are displayed in **Tables 1.28** and **1.29**.

In 2020, PUNR was 31.0 percent in samples A–Q with notable differences between the samples. L1 shows the best measures with 13.6 percent, while PUNR in sample N is at 36.4 percent. In part, these differences might be due to the different heritage of these two samples. Sample L1 was introduced to the households as a family study, stressing the importance of both parents' participating in the survey. Meanwhile, the source of households in sample N is the study PIACC (Programme for the International Assessment of Adult Competencies) which focused on individuals only in the first wave. The same is true for samples P and Q which had an anchor-respondent design in their respective first waves.

Table 1.28: **Partial unit non-response by sample (PUNR, in percent)**

	Total	A-H	J-K	L1	L2/3	N	O	P	Q
PUNR <sup>1</sup>	31.0	25.0	23.4	13.6	29.7	36.4	27.4	62.3	51.0

<sup>1</sup> Share of households (number of household members > 1) with at least one missing individual questionnaire.

Table 1.29: **Partial unit non-response<sup>1</sup> by type of fieldwork 2020 (PUNR, in percent)**

	Total	Interviewer-based	MAIL <sup>2</sup>	CAWI <sup>3</sup>
PUNR <sup>1</sup>	31.0	30.6	32.0	36.5

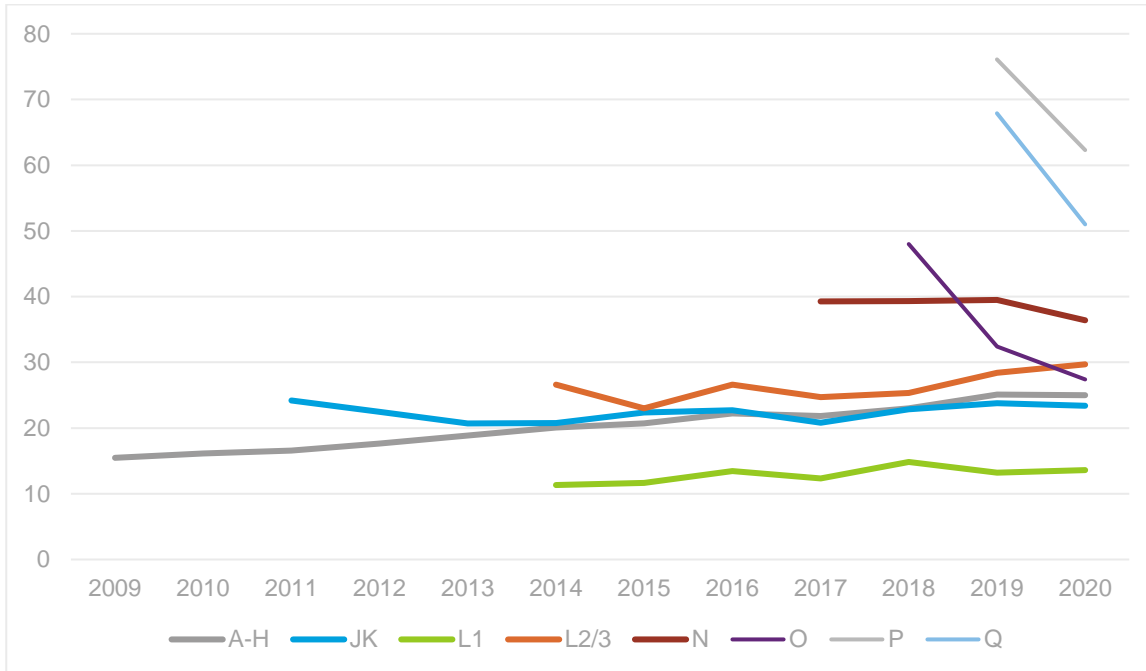
<sup>1</sup> Share of households (number of household members > 1) with at least one missing individual questionnaire.

<sup>2</sup> All households in the MAIL mode used to be dropouts in the interviewer-based mode but were willing to continue participation via mail.

<sup>3</sup> Households with start mode CAWI automatically switch to interviewer-based mode if they fail to participate online. Households from sample L2/3 with start mode interviewer can also switch to CAWI, if they express interest in participating online.

The results in **Table 1.29** show that there are also differences in PUNR for the modes. Interviewers seem to be slightly more successful in generating complete households, with a PUNR of 30.6 percent compared to MAIL or CAWI, with rates of 32 percent and 36.5 percent respectively.

Figure 1.5: Development partial unit non-response<sup>1</sup> since 2009 (PUNR, in percent)



<sup>1</sup> Share of households (number of household members > 1) with at least one missing individual questionnaire.

Looking at the last ten years, there is a slight upward trend visible concerning PUNR for most of the samples (**Figure 1.5**). This trend affects mainly the oldest samples A-H as well as sample L1, albeit at a generally much lower level of partial unit non-response. PUNR for sample N remains relatively high while there is a clear downward trend for sample O.

## 1.6.3 Participation by types of questionnaires

The numbers of completed questionnaires for each sample or sample group respectively are presented in **Table 1.30**. Most individual questionnaires were completed in sample group A-H, followed by J-K and N. The highest number of life-history questionnaires was answered in one of the newest SOEP-samples, sample P. In sample L2/3, high numbers of youth questionnaires were realized because of the specific age structure and high number of children in these samples. L1 stands out with a high number of questionnaires for mothers and older children as well as young teenagers because it consists of families with at least one child born in the years 2007 to 2010.

Table 1.30: Number of interviews by sample and questionnaire

	Total	A-H	J-K	L1	L2/3	N	O	P	Q
Household questionnaire	13,460	4,726	2,265	866	1,538	1,844	569	1,229	423
Individual questionnaire	21,615	7,541	3,588	1,590	2,856	2,941	814	1,739	546
Life-history questionnaire	747	110	57	37	174	104	27	189	49
Youth questionnaire: age 16 or 17	390	74	40	55	154	39	7	18	3
Cognitive competency tests	99	25	23	34	-	15	2	-	-
Youth questionnaire: age 13 or 14	359	52	36	79	118	34	15	21	4
Youth questionnaire: age 11 or 12	521	71	34	209	106	55	16	26	4
Mother and child questionnaire: age 9 or 10	472	58	27	248	39	48	13	35	4
Questionnaire for parents: age 7 or 8 <sup>1</sup>	471	109	64	133	38	62	19	39	7
Questionnaire for parents: age 7 or 8 <sup>2</sup>	291	70	41	75	25	35	14	26	5
Mother and child questionnaire: age 5 or 6	254	68	45	43	21	44	9	23	1
Mother and child questionnaire: age 2 or 3	230	73	29	24	21	38	20	23	2
Mother and child questionnaire: newborn	173	60	41	16	15	25	10	4	2
Questionnaire "Gap"	706	146	111	68	171	153	57	-	-
Questionnaire "Deceased person"	345	110	70	22	30	55	16	31	11

<sup>1</sup> Total number of questionnaires filled out by either mother or father.

<sup>2</sup> Number of children for which at least one questionnaire was filled out.

**Table 1.31** shows again that CAPI is the interviewing mode applied for the majority of interviews. In 2020, 52.7 percent of all individual questionnaires were completed via CAPI. Apart from CAPI, self-interviewing is the most widespread mode for individual interviews, with 6.591 cases in 2020. This is also true for youth questionnaires. The sharp increase in modes other than CAPI is due to the lifting of mode restrictions in 2020.

Table 1.31: Number of interviews by mode and questionnaire

		Interviewer-Based		Centrally Administered			Total <sup>1</sup>
		CAPI <sup>3</sup>	PAPI	SELF	MAIL	CAWI	
Household questionnaire	abs.	7,451	205	3,742	1,316	613	13,327
	In %	55.9	1.5	28.1	9.9	4.6	100.0
Individual questionnaire	abs.	11,208	288	6,591	2,026	1,168	21,281
	In %	52.7	1.4	31.0	9.5	5.5	100.0
Life-history questionnaire	abs.	345	15	282	33	63	738
	In %	46.7	2.0	38.2	4.5	8.5	100.0
Youth questionnaire: age 16 or 17	abs.	170	4	120	26	66	386
	In %	44.0	1.0	31.1	6.7	17.1	100.0
Cognitive competency tests	abs.	-	-	99	-	-	99
	In %	-	-	100.0	-	-	100.0
Youth questionnaire: age 13 or 14	abs.	154	5	132	17	46	354
	In %	43.5	1.4	37.3	4.8	13.0	100.0
Youth questionnaire: age 11 or 12	abs.	231	6	209	21	47	514
	In %	44.9	1.2	40.7	4.1	9.1	100.0
Mother and child questionnaire: age 9 or 10	abs.	221	6	202	22	17	468
	In %	47.2	1.3	43.2	4.7	3.6	100.0
Questionnaire for parents: age 7 or 8 <sup>2</sup>	abs.	198	10	208	34	14	464
	In %	42.7	2.2	44.8	7.3	3.0	100.0
Mother and child questionnaire: age 5 or 6	abs.	102	0	0	29	7	138
	In %	73.9	0.0	0.0	21.0	5.1	100.0
Mother and child questionnaire: age 2 or 3	abs.	105	0	0	31	6	142
	In %	73.9	0.0	0.0	21.8	4.2	100.0
Mother and child questionnaire: newborn	abs.	82	0	0	28	4	114
	In %	71.9	0.0	0.0	24.6	3.5	100.0
Questionnaire "Gap"	abs.	341	30	126	111	89	697
	In %	48.9	4.3	18.1	15.9	12.8	100.0
Questionnaire "Deceased person"	abs.	252	0	0	8	1	261
	In %	96.6	0.0	0.0	3.1	0.4	100.0

<sup>1</sup> For missing interviews, in comparison with the previous table, no information for mode was available.

<sup>2</sup> Total number of questionnaires filled out by either mother or father.

<sup>3</sup> Including CAPI-TEL



Response rates for the different questionnaires are reported in **Table 1.32**. These rates are calculated only for members of households that participated in the survey in the current wave. For the youth questionnaires, response rates are between 79.2 percent and 89.1 (age 11 and 12). For mother and child questionnaires, response rates well over 90 percent are usual with exception for the “Mother and child questionnaire: newborn”.

Table 1.32: **Response rates by questionnaire**

	Gross sample <sup>1</sup>	Number of interviews	Response Rate
Individual questionnaire	24,889	21,541	86.5
Life-history questionnaire	769	733	95.3
Youth questionnaire: age 16 or 17	491	389	79.2
Cognitive competency tests	210	99	47.1
Youth questionnaire: age 13 or 14	413	358	86.7
Youth questionnaire: age 11 or 12	585	521	89.1
Mother and child questionnaire: age 9 or 10	502	470	93.6
Questionnaire for parents: age 7 or 8 <sup>2</sup>	302	270	89.4
Mother and child questionnaire: age 5 or 6	276	254	92.0
Mother and child questionnaire: age 2 or 3	247	230	93.1
Mother and child questionnaire: newborn	208	172	82.7
Questionnaire “Gap”	893	706	79.1
Questionnaire “Deceased person”	170	70	41.2

<sup>1</sup> Gross sample = target population in participating households (one household questionnaire and at least one individual questionnaire available), without household members who are deceased or have moved abroad.

<sup>2</sup> Number of children for which at least one questionnaire was filled out.

## 1.6.4 Interview length per questionnaire

Figures for the median interview length for each questionnaire and the different interviewing modes respectively are presented in **Table 1.33**. The median interview length for the household questionnaire amounts to 15 minutes in total. The individual questionnaire took 40 minutes to complete, with 90 percent of the interviews lasting between 20 and 90 minutes.

Table 1.33: **Median interview length (minutes) by questionnaire and mode<sup>1</sup>**

	Total		CAPI	PAPI	SELF	CAWI
	Median	Percentiles (5%, 95%)	Median	Median	Median	Median
Household questionnaire	15	(10, 45)	15	20	25	20
Individual questionnaire	40	(20, 70)	40	40	45	30
Life-history questionnaire	30	(15, 69)	24	30	30	-
Youth questionnaire: age 16 or 17	30	(20, 60)	35	25	35	30
Cognitive competency tests	Not recorded					
Youth questionnaire: age 13 or 14	30	(15, 50)	25	25	30	25
Youth questionnaire: age 11 or 12	30	(15, 52)	25	28	30	25
Mother and child questionnaire: age 9 or 10	15	(10, 30)	15	25	15	15
Questionnaire for parents: age 7 or 8	10	(6, 25)	10	15	11	11
Mother and child questionnaire: age 5 or 6	Not recorded					
Mother and child questionnaire: age 2 or 3	Not recorded					
Mother and child questionnaire: newborn	Not recorded					
Questionnaire "Gap"	5	(2, 30)	5	5	5	-
Questionnaire "Deceased person"	Not recorded					

<sup>1</sup> For MAIL mode, interview length is not recorded.

## 1.6.5 Consent to record linkage

**Table 1.34** presents figures regarding the respondents' consent to linking their survey data to data from the German statutory pension insurance (Deutsche Rentenversicherung – DRV). In 2018, it was the first time that we tried to get this consent to record linkage with a consent rate of 49.7 percent. In 2020, we asked all the respondents where the consent was missing. Therefore, the consent rate in 2020 was 40.5 percent of respondents who were marked for record linkage. The numbers are slightly lower for the older samples A-H and considerably higher for all newer samples, especially in sample N. Most respondents gave their consent electronically.

Table 1.34: **Consent to record linkage**

	Total	A-H	J-K	L1	N	O
Gross sample <sup>1</sup>	10,755	2,996	1,896	831	2,890	1,318
Consent	4,358	879	622	308	1,609	462
Consent rate (in %)	40.5	29.3	32.8	37.1	55.7	35.1
Consent on paper	2,355	632	254	160	860	245
Consent in CAPI	2,003	247	368	148	749	217
Consent in CAPI (in %)	46.0	28.1	59.2	48.1	46.6	47.0

<sup>1</sup> Participating respondents in 2020, born in 2002 or earlier, in CAPI, CAWI and PAPI mode.

## 1.7 Data preparation

In a large-scale panel study such as the SOEP, data processing and data preparation are fundamentally important processes of quality management. During the entire process, we must ensure that we generate and provide optimal data quality that is consistent over the years: Only then can the necessary longitudinal consistency be generated which is required by the SOEP users that conduct panel analyses. Data processing and data preparation are elements of complex and multistage work processes. Moreover, the procedures and techniques of the data checking and data cleaning are constantly being further developed: partly through learning processes that are “content-related”, partly through optimizations that occur as a result of IT innovations. However, over the years, the main features remain unchanged and have been presented in more detail in the field report 2011 by Kantar<sup>5</sup>. The DIW receives the (net) data in two forms, the adjusted and unadjusted data sets. If required, data inferences can thus be identified and traced any time on a case-by-case basis.

The major elements of data processing and data preparation in the SOEP are data collection, data checking, data cleaning and data enrichment:

- **Data collection** includes all activities of the production process that serve to provide the collected data for further processing and preparation of the data.
- **Data checking** is the most important element of quality management, because the checking criteria defined here specify to what extent and at which points checks are carried out, and consequently, to what extent the collected raw data is edited.
- **Data cleaning** includes all direct inferences into the collected data at the individual case level, which are usually corrections of data errors in order to generate cross-sectional and panel consistency.
- In the **data enrichment** process, new information (both at the individual and the aggregated level) is generated and added to the data set as an additional variable. This means e.g. encoding open answers or adding spatial indicators.

The various elements of the data processing procedures cannot be strictly separated. They are rather interdependent and continuous processes. For instance, certain checking steps are already integrated into the data collection, and a large part of the data cleaning is conducted as part of the data checking. In addition to regarding the individual elements, it is useful to outline the process stages that emerge during data processing.

### Preparatory activities:

- Creating or revising the data-entry screens and check programs
- Programming the data entry screens (scan programs)
- Testing the programs

### Collection of data and address protocols; first preliminary checks:

- Scanning the paper questionnaires (including digital archiving of the scanned questionnaires)
- Converting all data collected with multimode into a uniform data format for data checking and delivery
- Collecting the address protocols (entry into the panel file)
- 1<sup>st</sup> checks on completeness of the households (correct quantity/type of completed questionnaires)
- First content checks (e.g. invalid multiple answers)
- Subsequent collection of certain information/missing questionnaires

---

<sup>5</sup> SOEP Methodenbericht 2011, Huber et al.

## **Gross related basic checking:**

- Checks on completeness and consistency of the various data sets per household (household data, individual data, additional questionnaire)
- Comparing the gross information (panel file) with questionnaire and address protocol information (identity, marital status, household composition, address)
- Clarifying inconsistencies/irregularities with the target households

## **Net related checking:**

- Automatic checking of the single questionnaires according to certain processing rules and automatic setting of codes or completion/cleaning of the data
- Key checking elements: filtering, invalid multiple responses, total checks, value range, implausible answers
- Manual case-by-case checks for unclear cases
- Panel checks (checking the updated activities calendar, checking status information and status changes)
- Editing the interviewers' notes and respondents' comments (e.g. omitting names and addresses)

## **Data enrichment/coding:**

- Coding open plaintext answers on the job title and the industry as well as the educational level obtained based on the most recent classification schemes
- Adding regional indicators (postcode, official municipality key, etc.)
- Adding Microm data (including geocoded addresses, GPS codes)
- Geocoding of places of birth inside and outside Germany

## **Final checking:**

- Checking the entire prepared data stock (household questionnaire, individual questionnaire, additional questionnaire, gross data sets) for completeness and possible irregularities (including the correction of program and checking errors)

## **Data delivery:**

- Transfer of pre-checked intermediate data after half of the field time
- Transfer of the code book (description of the variables and documentation of the coding rules)
- Transfer of the final integrated and standardized gross and net data files (anonymized) in checked and unchecked form

It should be mentioned here that the individual process steps are not conducted in a fixed order, but mostly simultaneously, interdependently and in several loops. In order to be able to deliver the checked, cleansed and enriched data promptly, the major part of the described work steps is organized "on a weekly basis" (according to "field weeks"). This means that the collected data are not gathered and processed at once at the end of the survey phase, but continuously right from the beginning of fieldwork. This means that almost all process steps are conducted on an ongoing basis.

## 1.8 Delivered data

### Gross Data

Gross data Household	hbru_A-Q_2020.sav, hbru_A-Q_2020.dta
Gross data Individuals	pbru_A-Q_2020.sav, pbru_A-Q_2020.dta
Interviewer data	intband_A-Q_2020.sav, intband_A-Q_2020.dta
SOEP Individuals A-Q	p_A-Q_2020.sav, p_A-Q_2020.dta
Qualitative Special Surveys 2013-2020	quali_A-Q_2020.sav, quali_A-Q_2020.dta

### Net Data

Net data Household checked	H20_A-Q.sav, H20_A-Q.dta
Net data Household unchecked	H20_A-Q_u.sav, H20_A-Q_u.dta
Net data Individuals checked	P20_A-Q.sav, P20_A-Q.dta
Net data Individuals unchecked	P20_A-Q_u.sav, P20_A-Q_u.dta
Net data Life history checked	L20_A-Q.sav, L20_A-Q.dta
Net data Life history unchecked	L20_A-Q_u.sav, L20_A-Q_u.dta
Net data Youth (age 16 or 17) checked	J20_A-Q.sav, J20_A-Q.dta
Net data Cognitive competency test	DJ20_A-Q.sav, DJ20_A-Q.dta
Net data Youth (age 13 or 14)	FJ20_A-Q.sav, FJ20_A-Q.dta
Net data Youth (age 11 or 12)	S20_A-Q.sav, S20_A-Q.dta
Net data Mother and child E (age 9 or 10)	ME20_A-Q.sav, ME20_A-Q.dta
Net data Questionnaire for parents (age 7 or 8)	EL20_A-Q.sav, EL20_A-Q.dta
Net data Mother and child C (age 5 or 6)	MC20_A-Q.sav, MC20_A-Q.dta
Net data Mother and child B (age 2 or 3)	MB20_A-Q.sav, MB20_A-Q.dta
Net data Mother and child A (newborn)	MA20_A-Q.sav, MA20_A-Q.dta
„Gap“ data from 2019	LUE19_20_A-Q.sav, LUE19_20_A-Q.dta
Net data “Deceased person”	VP20_A-Q.sav, VP20_A-Q.dta

### Other Data

Professions, sectors, final coding	BerufeBranchenAusbildung_2020_A-Q.dta
University coding	Hochschul-Abschl_2020_A-Q.dta

## 2 Samples M1/2

### 2.1 Introduction

**Table 2.1** summarizes the main characteristics of the 2020 wave in samples M1/2. 1,266 households took part in the study between April and August. This results in an adjusted response rate of 76.8 percent, which is a slight increase compared to the wave of 2019 (72.6 percent). Participation of households that were respondents in the previous wave was at 82.8 percent. In all households, 2,292 adults and 153 youths gave interviews. For an additional 349 children of various age groups, data are available from interviews with a parent. 13 different questionnaires were fielded in up to four different modes. Partial unit non-response (PUNR<sup>6</sup>) was at 29.4 percent, a slightly lower value than in the previous wave (31.1 percent).

Table 2.1: **Summary fieldwork**

Fieldwork period	April - August
Mode (main questionnaires)	CAPI, PAPI, SELF
Gross sample (hh)	1,720
Net sample (hh)	1,296
Response rate (adjusted; hh) <sup>1</sup>	Overall: 76.8 Previous wave respondents: 82.8 Previous wave dropouts: 44.5 New households: 61.1
Number of questionnaires	Adults: 4 Youths: 3 Children: 5
Net sample (individuals)	Adults: 2,292 Youths: 153 Children: 349
Questionnaire length (median, in minutes)	Household: 15 Adult <sup>2</sup> : 40
Partial unit non response	29.4

<sup>1</sup> RR = percentage of all households with at least one hh and individual interview in the gross sample (gross sample adjusted for households where the last person is deceased or the household moved abroad, is permanently untraceable or dissolved households where the last member moved into another SOEP household).

<sup>2</sup> Without first-time respondents.

<sup>6</sup> PUNR: share of households (number of household members > 1) with at least one missing individual questionnaire.

## 2.2 Development samples M1/2

The two subsamples that constitute the SOEP Migration Survey, which was designed to improve the representation of migrants living in Germany, are samples M1 and M2. Established in 2013, 2,723 households with at least one person with a migration background were interviewed to enhance the analytic potential for integration research and migration dynamics. A second migration sample (Sample M2) of 1,096 households was added in 2015.

Figure 2.1: Overview of migration samples

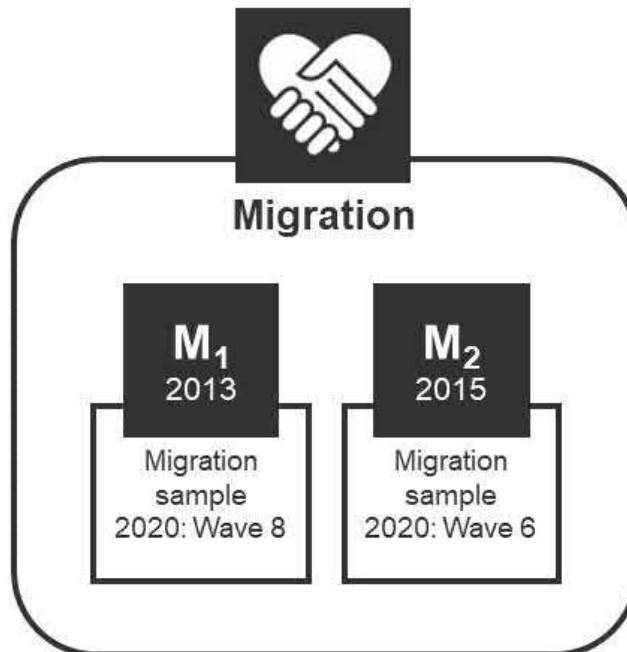
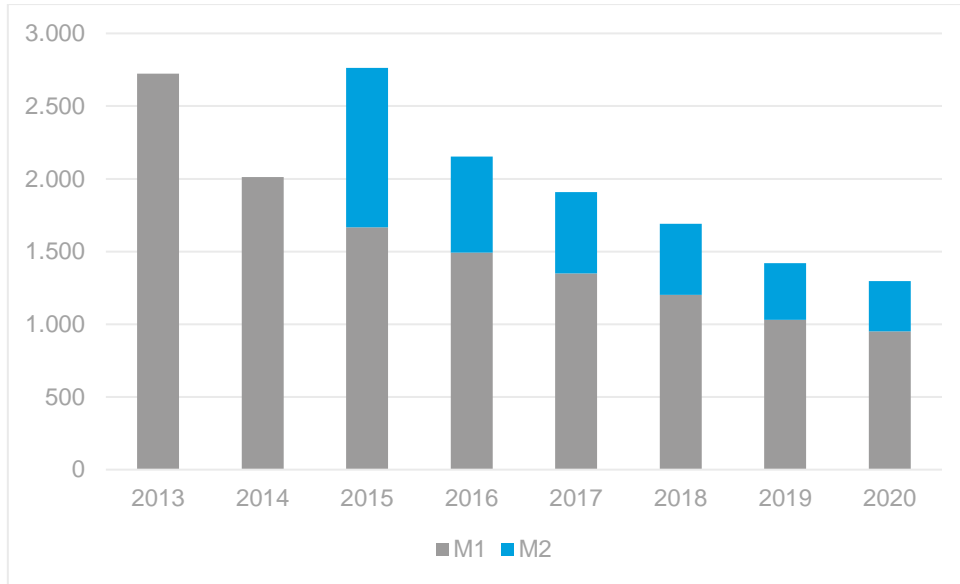




Figure 2.2: Development of samples since 2013 (number of households)



## 2.3 Structure of the gross sample

As presented in **Table 2.2**, 2,015 households were part of the gross sample of M1/2 in 2019, which is the basis for the administered gross sample for 2020. 70.5 percent of these households successfully participated without refusal for the next survey of 2020. 590 households (29.3 percent) did not participate in the survey in 2019. 250 households only dropped out temporarily and were part of this year's gross sample. 231 of the dropouts were final, and another 109 households were excluded for dropping out temporarily in two consecutive waves. Comparing the two samples M1 and M2 reveals that dropout rates were higher in the more recent sample M2.

Table 2.2: Final gross sample 2019 by type of household

Household level	Total		M1		M2	
	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample</b>	2,015	100.0	1,424	100.0	591	100.0
Participating households without refusal for next wave	1,420	70.5	1,029	72.3	391	66.2
Participating households with refusal for next wave	1	0.0	1	0.1	0	0.0
Temporary dropouts	250	12.4	170	11.9	84	14.2
Final dropouts	231	11.5	148	10.4	83	14.0
2-year rule	109	5.4	76	5.3	33	5.6

Table 2.3: Administered gross sample 2020 by type of household

Household level	Total		M1		M2	
	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample</b>	1,720	100.0	1,234	100.0	486	100.0
Participating households in previous wave	1,421	82.6	1,030	83.5	391	80.5
Temp. dropout in previous wave	254	14.8	170	13.8	84	17.3
New households (Split-off hhs)	45	2.6	34	2.8	11	2.3

**Table 2.3** presents total and sample-related figures for the administered gross sample in 2020 distinguished by three different types of households. A total of 1,720 households was processed in the two samples M1 and M2. 82.6 percent of these households participated in the previous wave. 14.8 percent dropped out temporarily in 2019 while 2.6 percent are new households that emerge because panel members move out of existing panel households.

Table 2.4: Household characteristics by sub-samples I

Household Level	Total		M1		M2	
	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample<sup>1</sup></b>	1.671	100.0	1,196	100.0	475	100.0
<b>HH size</b>						
1	286	17.1	198	16.6	88	18.5
2	440	26.3	334	27.9	106	22.3
3	366	21.9	247	20.7	119	25.1
4	360	21.5	265	22.2	95	20.0
5+	219	13.1	152	12.7	67	14.1
<b>State</b>						
Schleswig-Holstein	40	2.4	33	2.8	7	1.5
Hamburg	56	3.4	46	3.8	10	2.1
Lower Saxony	154	9.2	110	9.2	44	9.3
Bremen	17	1.0	12	1.0	5	1.1
North Rhine-Westphalia	463	27.7	351	29.3	112	23.6
Hesse	118	7.1	79	6.6	39	8.2
Rhineland Palatinate	112	6.7	73	6.1	39	8.2
Saarland	15	0.9	14	1.2	1	0.2
Baden-Wuerttemberg	207	12.4	150	12.5	57	12.0
Bavaria	287	17.2	188	15.7	99	20.8
Berlin	64	3.8	39	3.3	25	5.3
Brandenburg	36	2.2	25	2.1	11	2.3
Mecklenburg Western Pomerania	7	0.4	7	0.6	0	0.0
Saxony	35	2.1	31	2.6	4	0.8
Saxony-Anhalt	27	1.6	24	2.0	3	0.6
Thuringia	33	2.0	14	1.2	19	4.0

<sup>1</sup> Status as reported at the end of wave 2019. So new households and households that rejoined the panel in 2020 but were not part of fieldwork 2019 are missing.

Figures for basic household characteristics are presented in **Table 2.4** and **Table 2.5**. About half of the households in samples M1 and M2 consist of two or three household members. However, households with a higher number of members are not uncommon. A higher share of households in samples M1/2 is located in the center of larger cities (BIK-type = 0 or 2) compared to samples A-O.

Table 2.5: Household characteristics by sub-samples (II)

Household Level	Total		M1		M2	
	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample<sup>1</sup></b>	1,675	100.0	1,200	100.0	475	100.0
<b>BIK type<sup>2</sup></b>						
0	577	34.4	400	33.3	177	37.3
1	163	9.7	119	9.9	44	9.3
2	332	19.8	238	19.8	94	19.8
3	148	8.8	116	9.7	32	6.7
4	19	1.1	11	0.9	8	1.7
5	127	7.6	109	9.1	18	3.8
6	181	10.8	117	9.8	64	13.5
7	109	6.5	78	6.5	31	6.5
8	11	0.7	7	0.6	4	0.8
9	8	0.5	5	0.4	3	0.6
<b>Community size<sup>3</sup></b>						
1	25	1.5	15	1.3	10	2.1
2	73	4.4	44	3.7	29	6.1
3	367	22.0	271	22.7	96	20.2
4	344	20.6	232	19.4	112	23.6
5	141	8.4	105	8.8	36	7.6
6	312	18.6	241	20.1	71	14.9
7	413	24.7	292	24.3	121	25.5

<sup>1</sup> Status as reported at the end of wave 2019. So new households and households that rejoined the panel in 2020 but were not part of fieldwork 2019 are missing.

<sup>2</sup> BIK-type: 0 (more than 500,000 inhabitants/center) 1 (more than 500,000 inh./periphery), 2 (100,000 to 499,999 inh./center)

3 (100,000 to 499,999 inh./periphery), 4 (50,000 to 99,999 inh./center), 5 (50,000 to 99,999 inh./periphery), 6 (20,000 to 49,999 inh.), 7 (5,000 to 19,999 inh.), 8 (2,000 to 4,999 inh.), 9 (fewer than 2,000 inh.)

<sup>3</sup> Community size: 1 (fewer than 2000 inhabitants), 2 (2,000 to 5,000 inh.), 3 (5,000 to 20,000 inh.), 4 (20,000 to 50,000 inh.), 5 (50,000 to 100,000 inh.), 6 (100,000 to 500,000 inh.), 7 (more than 500,000 inh.).

## 2.4 Questionnaires and fieldwork material

### 2.4.1 Questionnaires

**Table 2.6** displays which questionnaires were available in which modes for samples M1 and M2. For the main questionnaires, household and individual, only interviews in CAPI mode are permitted. But most supplementary questionnaires were also offered in PAPI mode. Only the cognitive competency test for 16- and 17-year-olds is provided solely as a paper questionnaire.

Table 2.6: **Questionnaires and modes**

	CAPI	PAPI
Household questionnaire	✓	-
Individual questionnaire (+ life-history module)	✓	-
Youth questionnaire: age 16 or 17	✓	✓
Cognitive competency test	-	✓
Youth questionnaire: age 13 or 14	✓	✓
Youth questionnaire: age 11 or 12	✓	✓
Mother and child questionnaire: age 9 or 10	✓	✓
Questionnaire for parents: age 7 or 8	✓	✓
Mother and child questionnaire: age 5 or 6	✓	✓
Mother and child questionnaire: age 2 or 3	✓	✓
Mother and child questionnaire: newborn	✓	✓
Questionnaire "Gap"	✓	-
Questionnaire "Deceased person"	✓	-

For data collection in the SOEP migration samples in 2020, all the questionnaires from SOEP-Core were used. The life-history questionnaire was adapted in samples A-Q to better capture the individual migration background. Those changes were carried over to samples M1 and M2, so that the life-history questionnaire is identical for all SOEP-CORE samples (with the exception of the refugee sub-population in samples M3-5). This instrument was used for adult household members who were participating in the study for the first time.

All new respondents with the individual questionnaire and the life history module got additionally the CAMCES module. This module recorded all educational achievements in the foreign language, who were acquired abroad.

## 2.4.2 Fieldwork material

In addition to the questionnaires, a whole range of fieldwork materials, such as letters, leaflets or documents for the interviewers, were designed, printed and sent to households and interviewers. **Table 2.7** provides an overview of the different material types that are prepared in samples M1/2. Because the mode was restricted to CAPI in these samples, the number of different versions of materials was notably smaller than in samples A-H. But many materials were provided in six different languages (German, English, Russian, Turkish, Romanian and Polish).

### **Advance letter**

About two weeks before the start of the fieldwork period, the households received an advance letter in which the interviewer's visit was announced.

### **Thank-you letter**

Sometime after participating in the study, each household received a thank-you letter with a postage stamp enclosed as a small gift.

### **Leaflet**

Every household in samples M1/2 received an eight-page leaflet with SOEP reports and published results. The leaflet was handed out by the interviewer.

### **Declaration on data protection**

Every household got access to a two-page declaration on data protection detailing the organizations responsible for processing all respondent data along with a description of data handling and data recipients. The sheet was also handed out by the interviewers.

### **Consent to record linkage form**

To all respondents born in 2002 or earlier and who have not been asked for their consent in one of the previous waves, a consent to record linkage form was presented that allows an individual linkage between the respondent's data and employment history data available at the Institute for Employment Research (Institut für Arbeitsmarkt und Berufsforschung, IAB Nuremberg). All respondents were also asked their consent to the linkage of their survey data to data from the German statutory pension insurance (Deutsche Rentenversicherung – DRV).

### **Address form and household grid**

The address form provided an overview of the household composition as it was last known to Kantar. The interviewers had to document when and with whom the survey was conducted or why a sample member did not participate in the current year. They were also asked to note every single contact attempt made. For samples M1/2, interviewers did this electronically in the "Mein Kantar" software.

Moreover, the interviewers were asked to carefully document any moves of households and household members or changes in household composition. An additional form had to be completed for every person that had left the household since the last survey because all of these remain potential respondents and were tracked even if the new address was unknown.

### **Showcards**

Showcards were used by interviewers during the CAPI interview to show respondents longer item lists or scales.

Table 2.7: Fieldwork material

CAPI	
Advance letter	<ul style="list-style-type: none"> <li>• In 6 languages (German, English, Russian, Turkish, Romanian, and Polish)</li> </ul>
Thank-you letter	<ul style="list-style-type: none"> <li>• In German language</li> </ul>
Leaflet	<ul style="list-style-type: none"> <li>• In 6 languages</li> </ul>
Declaration on data protection	<ul style="list-style-type: none"> <li>• In 6 languages</li> </ul>
Consent to record linkage form <sup>1</sup>	<ul style="list-style-type: none"> <li>• In German language</li> </ul>
Address form and household grid <sup>2</sup>	<ul style="list-style-type: none"> <li>• Electronic form (Mein Kantar)</li> </ul>
Showcards	<ul style="list-style-type: none"> <li>• In German language</li> </ul>
Project instruction book	<ul style="list-style-type: none"> <li>• Version for M1-2</li> </ul>
Other interviewer material	<ul style="list-style-type: none"> <li>• Project description</li> <li>• Contact card</li> <li>• HH information card</li> </ul>

<sup>1</sup> Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB Nuremberg): Link to employment history data and German statutory pension insurance (Deutsche Rentenversicherung – DRV)

<sup>2</sup> Including the so-called “B3 Form” used to process address changes.

### Project instruction book

In addition to a shorter project description, interviewers in samples M1/2 also received a much more detailed instruction manual that is about 60 pages long. This manual contained information on special features of the current wave, specific processing instructions and questionnaires as well as background information on the project.

### Other interviewer material

In addition, the interviewers received contact cards for households that could not be reached at home and household information cards with information on individual numbers, names, years of birth, types of questionnaires, incentives, survey modes, and notes.

## 2.5 Conducting the survey

### 2.5.1 Survey mode

The mode for the main questionnaires in samples M1/2 (household, individual) is usually restricted to CAPI, which resulted in a 100% share of CAPI interviews for these interviews. However, due to the COVID-19 pandemic, these usually rigorous mode restrictions were lifted to allow for more flexibility. This resulted in an unusually high number of SELF and PAPI interviews in samples, M1 and M2. In addition to the PAPI and SELF options being opened, interviews could also be conducted via phone using the CAPI-instrument. **Tables 2.8a and 2.8b** display the distribution of modes across samples M1 and M2 on the individual and on the household level.

Table 2.8a: Interviewing modes by sub-samples – individual questionnaire

Individual level		Interviewer-Based				Total
		CAPI-TEL	CAPI	PAPI	SELF	
M1	Abs.	564	427	22	659	1.672
	In %	33.7	25.5	1.3	39.4	100.0
M2	Abs.	142	211	0	201	554
	In %	25.6	38.1	0.0	36.3	100.0
Total	Abs.	706	638	22	860	2.226
	In %	31.7	28.7	1.0	38.6	100.0

Table 2.8b: Interviewing modes by sub-samples – household questionnaire

Household level		Interviewer-Based				Total
		CAPI-TEL	CAPI	PAPI	SELF	
M1	Abs.	331	252	6	340	929
	In %	35.6	27.1	0.6	36.6	100.0
M2	Abs.	96	125	0	116	337
	In %	28.5	37.1	0.0	34.4	100.0
Total	Abs.	427	377	6	456	1.266
	In %	33.7	29.8	0.5	36.0	100.0

An overview of the different interviewing modes used for the youth questionnaires is provided in **Table 2.8c**. The standard CAPI and SELF interviewing were the dominant modes with 47.0 percent and 46.4 percent respectively. The CAPI-TEL option was used in six cases only across both samples.



**Table 2.8c: Interviewing modes by sub-samples – youth questionnaires**

Individual level	Interviewer-Based					Total
	CAPI-TEL	CAPI	PAPI	SELF		
M1	Abs.	6	57	4	60	127
	In %	4.7	44.9	3.1	47.2	100.0
M2	Abs.	0	14	0	10	24
	In %	0.0	58.3	0.0	41.7	100.0
Total	Abs.	6	71	4	70	151
	In %	4.0	47.0	2.6	46.4	100.0

As shown in **Table 2.9**, with 67.5 percent, the proportion of interviews conducted in CAPI mode is even higher for the child questionnaires. Self-interviewing was done in 24.4 percent of all cases. CAPI via phone was used in only 6.4 percent of all cases.

**Table 2.9: Interviewing modes by sub-samples – child questionnaires**

Individual level	Interviewer-Based					Total
	CAPI-TEL	CAPI	PAPI	SELF		
M1	Abs.	9	101	3	41	154
	In %	5.8	65.6	1.9	26.6	100.0
M2	Abs.	6	57	1	16	80
	In %	7.5	71.3	1.3	20.0	100.0
Total	Abs.	15	158	4	57	234
	In %	6.4	67.5	1.7	24.4	100.0

## 2.5.2 Fieldwork timing

The fieldwork started in April and ended in September. **Table 2.10** shows figures for the fieldwork progress referring to the gross sample as well as the net sample. The most productive months in terms of conducted interviews were April, May and June. The following months were mainly used to convert soft refusals and access difficult-to-reach households.

Table 2.10: **Monthly fieldwork progress**

Household level	Gross Sample		Net Sample	
	Abs.	In %	Abs.	In %
April <sup>1</sup>	832	48.4	734	56.6
May	334	19.4	277	21.4
June	126	7.3	83	6.4
July	393	22.8	188	14.5
August	35	2.0	14	1.1
September	0	0.0	0	0.0

<sup>1</sup> Including households that refused to take part in the survey prior to start of fieldwork.

## 2.5.3 Translations

**Table 2.11** provides an overview of the availability of questionnaires in different languages. In 2020, the household questionnaire as well as the individual and life-history questionnaires were available in five alternative languages other than German, namely English, Russian, Turkish, Romanian and Polish.

Table 2.11: **Translated paper questionnaires**

	English	Russian	Turkish	Romanian	Polish
Household questionnaire	✓	✓	✓	✓	✓
Individual questionnaire	✓	✓	✓	✓	✓
Life-history questionnaire	✓	✓	✓	✓	✓
Youth questionnaires <sup>1</sup>	-	-	-	-	-
Mother and child questionnaires/ questionnaire for parents	-	-	-	-	-
Additional questionnaires <sup>2</sup>	-	-	-	-	-

<sup>1</sup> Including cognitive competency test.

<sup>2</sup> Gap, Deceased person.

The following **Table 2.12** contains information about language problems and the usage of translated questionnaires in samples M1 and M2. Of all conducted interviews, 17.0 percent required assistance with language problems. For those interviews, translated paper questionnaires were used in most cases. Also, assistance by a German-speaking person in the same household was relatively frequent.

**Table 2.12: Language problems and usage of translated paper questionnaires**

	Total	In % Net Sample
Net sample (individual questionnaire) <sup>1</sup>	2,292	100.0
No language problems occurred/no need for assistance with language problems	1,902	83.0
Assistance with language problems needed	390	17.0
<b>Of that number:</b>		
German-speaking person in the same household <sup>2</sup>	170	7.4
German-speaking person from outside the household <sup>3</sup>	35	1.5
Professional interpreter	5	0.2
Translated paper questionnaire <sup>4</sup>	144	6.2
<b>Of that number:</b>		
English	20	0.9
Russian	65	2.8
Turkish	10	0.4
Romanian	28	1.2
Polish	21	0.9

1 Including all individual questionnaires even if the household in which they are conducted is classified as a non-participating household.

2 In 11 cases both a German-speaking person in the same household and a translated paper questionnaire were used for assistance.

3 In 4 cases both a German-speaking person from outside the household and a translated paper questionnaire were used for assistance.

4 In 1 case both a Russian and Romanian translated paper questionnaire were used for assistance.

## 2.5.4 Panel maintenance and incentives

The incentives for respondents in samples M1/2 were very similar to those in the newer samples J, K and L1 within samples A-O. Adult participants received cash incentives from the interviewers after they completed their interview. The individual questionnaire was rewarded with 10 euros, while the shorter household questionnaire was rewarded with 5 euros. Children in the households that participated in one of the youth

questionnaires were presented with a small gift such as a power bank to charge their phone or a bicycle repair kit (**Table 2.13**).

In addition to the individual incentives, interviewers brought a small gift to all households which was presented upon arrival. This year's gift was a high-quality shopping bag.

Table 2.13: **Incentives**

M1/2	
Incentives for adults	HH: 5 euros Adult: 10 euros
Incentives for youth and child questionnaire	Youth questionnaire: age 16 or 17: Power bank Youth questionnaire: age 13 or 14: USB-Stick Youth questionnaire: age 11 or 12: Bicycle repair kit Child questionnaires: No incentive

## 2.5.5 Movers and tracing

Figures on movers and sources of new addresses are given in **Table 2.14**. Of 1,720 households in the gross sample in 2020, 12.3 percent were movers or new households. The tracing of those households was successful in 90.1 percent of all cases. The most important sources of information about addresses were the interviewers (44.8 percent), followed by the local registration offices (13.2 percent) that provide information about residents' current addresses if there is a scientific interest behind the inquiry. Another significant source is the postal service that movers sometimes notify of their new address (16.0 percent).

Table 2.14: **Movers and sources of new addresses of administered sample 2020**

Household level	Total	
	Abs.	In %
Gross sample	1,720	100.0
Movers and new households	212	12.3
Tracing success		
Tracing successful	191	90.1
Tracing not successful	21	9.9
Source		
Interviewer	95	44.8
Postal service	34	16.0
Local registration offices	28	13.2
Participant	5	2.4
Post-Address Research	50	23.6

## 2.5.6 Interviewer characteristics, training & monitoring

Information on some interviewer characteristics is collected in **Table 2.15**. In total, 170 interviewers were involved in samples M1 and M2 in 2020. At 61.8 percent, the share of male interviewers is significantly higher than the share of female ones. The majority of interviewers are between 60 and 79 years old and are responsible for fewer than 20 households in samples M1/2.

Table 2.15: Interviewer characteristics

Interviewer level	Total		M1		M2	
	Abs.	In %	Abs.	In %	Abs.	In %
<b>Number of interviewers</b>	170	100,0	154	100,0	106	100,0
<b>Gender</b>						
Male	105	61.8	97	63.0	62	58.5
Female	65	38,2	57	37.0	44	41.5
<b>Age</b>						
21-39 years	8	4.7	7	4.5	2	1.9
40-59	35	20.6	33	21.4	22	20.8
60-79	114	67.1	102	66.2	72	67.9
80+	13	7.6	12	7.8	10	9.4
<b>Number of households (gross)</b>						
Fewer than 5	61	34.3	67	41.6	71	62.8
5 – 19	100	56.2	84	52.2	41	36.3
20 – 99	17	9.6	10	6.2	1	0.9
More than 99	0	0.0	0	0.0	0	0.0

Interviewer experience in samples M1/2 is comparable to the newer samples within samples A-O. Many interviewers fall into the category of 5-9 years of experience with the SOEP while having a bit more experience with working with Kantar in general (**Table 2.16**).

Interviewer training for samples M1/2 usually takes part together with the training for samples A-O at the beginning of the year. This was also the case in 2020 (see **Section 1.5.6**).

Table 2.16: Interviewer experience

Interviewer level	Total		M1		M2	
	Abs.	In %	Abs.	In %	Abs.	In %
<b>Number of Interviewers</b>	170	100.0	154	100.0	106	100.0
<b>Experience with Kantar</b>						
0-4 years of experience with Kantar	29	17.1	28	18.2	14	13.2
5-9 years	45	26.5	40	26.0	26	24.5
10-19 years	71	41.8	63	40.9	45	42.5
20-29 years	15	8.8	14	9.1	13	12.3
30-39 years	5	2.9	4	2.6	3	2.8
More than 40 years	5	2.9	5	3.2	5	4.7
<b>Experience with SOEP</b>						
0-4 years	35	20.6	33	21.4	19	17.9
5-9 years	69	40.6	63	40.9	40	37.7
10-19 years	48	28.2	41	26.6	31	29.2
20-29 years	13	7.6	12	7.8	11	10.4
30-34 years	2	1.2	2	1.3	2	1.9
More than 35 years	3	1.8	3	1.9	3	2.8

Kantar places high priority on interviewer monitoring and has put an ISO-certificated process in place that is audited regularly. Kantar adheres to the German Business Association of Market and Social Research Institutes (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V., ADM) standards for internal regulation and monitoring of all systems and procedures. This means that a minimum of 10 percent of Kantar's annual interviews are checked, and every interviewer is monitored at least once a year.

In 2018, we expanded these existing quality-control measures for interviewer monitoring in the SOEP projects. The basic interviewer monitoring concept was expanded to a project-based control system for all SOEP samples: A number of participating households was contacted shortly after the interview by letter or phone asking them to confirm their participation in a regularly conducted interview. In case of inconsistencies and/or irregularities, we attempted to gain clarification through direct contact with respondents, primarily by telephone. We have also begun building a comprehensive suite of bespoke analytical tools to uncover irregularities in the survey data and para data. The indicators produced by this tool are can led to further measures taken in the re-contacting process.

Table 2.17: Interviewer monitoring

	Total		
	Abs.	In % (net sample)	In % (re-contacted household)
By households			
Households in net sample	1,296	100.0	-
Re-contacted households	676	52.2	100.0
Households with feedback	242	18.7	35.8
By interviewers			
Interviewers in net sample	173	100.0	-
Interviewers with re-contacted households <sup>1</sup>	162	93.6	-
Interviewers with non-standard behavior	0	0.0	-

<sup>1</sup> Only interviewers with few net cases are missing from the re-contacts.

**Table 2.17** shows details for the re-contacting process in samples M1/2. In existing panel samples, the aim is to re-contact at least 10 percent of households for every interviewer in every SOEP project. In M1/2, there is an additional requirement to re-contact at least 30 percent of households for every interviewer who is responsible for more than 1.5 percent of the households in the net samples. Hence, 52.2 percent of households in samples M1/2 were re-contacted. For these households whose interviewers had a response rate significantly below 30 percent, a short paper questionnaire was administered, which was then followed up by phone.

Overall, we re-contacted 676 households in samples M1/2 and received feedback from 242 households. This results in a response rate of 35.8 percent overall. We did not identify any interviewers who had not adhered to our standards in conducting interviews in the 2020 fieldwork period.



## 2.6 Fieldwork results

### 2.6.1 Participation on household level

Numbers of participation for three different types of households are provided in **Table 2.18**. Of 1,720 households in the gross sample, 1,296 households were successfully interviewed, which amounts to 75.3 percent.

Table 2.18: **Participation by type of household (with AAPOR codes)**

	Total		Respondents in previous wave		Dropouts in previous wave		New households	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
Gross sample	1,720	100.0	1,421	100.0	254	100.0	45	100.0
<b>Interview (1.0)</b>	1,296	75.3	1,165	82.0	109	42.9	22	48.9
Complete (1.1)	998	58.0	907	63.8	75	29.5	16	35.6
Partial (1.2)	298	17.3	258	18.2	34	13.4	6	13.3
<b>Non-interview (2.0)</b>	424	24.7	256	18.0	145	57.1	23	51.1
<b>Temporary dropout</b>	307	17.8	198	13.9	103	40.6	6	13.3
Non-contact (2.20)	115	6.7	63	4.4	51	20.1	1	2.2
Temporarily physically or mentally unable/incompetent (2.321)	1	0.1	1	0.1	0	0.0	0	0.0
Temporary refusal (2.351; 2.353; 2.354)	180	10.5	130	9.1	48	18.9	2	4.4
Household could not be traced (temporary) (3.18; 2.4)	7	0.4	3	0.2	1	0.4	3	6.7
Other temp. (2.52)	4	0.2	1	0.1	3	1.2	0	0.0
<b>Final dropout</b>	117	6.8	58	4.1	42	16.5	17	37.8
Permanent refusal (2.111)	85	4.9	44	3.1	33	13.0	8	17.8
Deceased (2.31)	0	0.0	0	0.0	0	0.0	0	0.0
Permanently physically or mentally unable (2.322)	0	0.0	0	0.0	0	0.0	0	0.0
Moved abroad (4.2)	13	0.8	10	0.7	3	1.2	0	0.0
Household dissolved (4.3)	3	0.2	1	0.1	2	0.8	0	0.0
Household untraceable (4.4)	16	0.9	3	0.2	4	1.6	9	20.0
Dropped out temporarily in two consecutive waves	0	0.0	0	0.0	0	0.0	0	0.0

Of the remaining 424 households, 307 temporarily dropped out of the survey (17.8 percent of the gross sample). Many of these households stated that they were currently unwilling to participate, e.g. due to lack of time. 117 households dropped out of the survey permanently (6.8 percent of the gross sample). Of all the households that temporarily dropped out in the previous wave, 42.9 percent participated again in 2020. Additionally, 48.9 percent of all new households were successfully integrated into the survey.

**Table 2.19** presents numbers of participation for both samples M1 and M2. At 77.1 percent, the participation rate in 2020 on household level was slightly higher for the larger sample M1. Temporary dropouts and permanent ones both occurred more often in M2.

Table 2.19: **Participation by sample (with AAPOR codes)**

	Total		M1		M2	
	Abs.	In %	Abs.	In %	Abs.	In %
Gross sample	1,720	100.0	1,234	100.0	486	100.0
<b>Interview (1.0)</b>	1,296	75.3	952	77.1	344	70.8
Complete (1.1)	998	58.0	735	59.6	263	54.1
Partial (1.2)	298	17.3	217	17.6	81	16.7
<b>Non-interview (2.0)</b>	424	24.7	282	22.9	142	29.2
<b>Temporary dropout</b>	307	17.8	207	16.8	100	20.6
Non-contact (2.20)	115	6.7	80	6.5	35	7.2
Temporary refusal (2.351; 2.353; 2.354)	180	10.5	116	9.4	64	13.2
Temporarily physically or mentally unable/incompetent (2.321)	1	0.1	1	0.1	0	0.0
Household could not be traced (temporary) (3.18; 2.4)	7	0.4	6	0.5	1	0.2
Other temp. (2.52)	4	0.2	4	0.3	0	0.0
<b>Final dropout</b>	117	6.8	75	6.1	42	8.6
Permanent refusal (2.111)	85	4.9	60	4.9	25	5.1
Permanently physically or mentally unable/incompetent (2.322)	0	0.0	0	0.0	0	0.0
Deceased (2.31)	0	0.0	0	0.0	0	0.0
Moved abroad (4.2)	13	0.8	3	0.2	10	2.1
Household dissolved (4.3)	3	0.2	2	0.2	1	0.2
Household untraceable (4.4)	16	0.9	10	0.8	6	1.2
Dropped out temporarily in two consecutive waves	0	0.0	0	0.0	0	0.0

**Table 2.20** provides response rates for three different types of households. In total, an adjusted response rate of 76.8 was achieved on household level in 2020. For households which responded in the previous wave, the rate was higher at 82.8 percent. Dropouts in the previous wave and new households usually attain notably lower response rates.

Table 2.20: **Response rate by type of household (in %)**

	Total	Respondents in previous wave	Dropouts in previous wave	New households
Response rate <sup>1</sup>	76.8	82.8	44.5	61.1

<sup>1</sup> RR = percentage of all households with at least one hh and individual interview in the gross sample (gross sample adjusted for households where the last person is deceased or the household moved abroad, is permanently untraceable or dissolved households where the last member moved into another SOEP household).

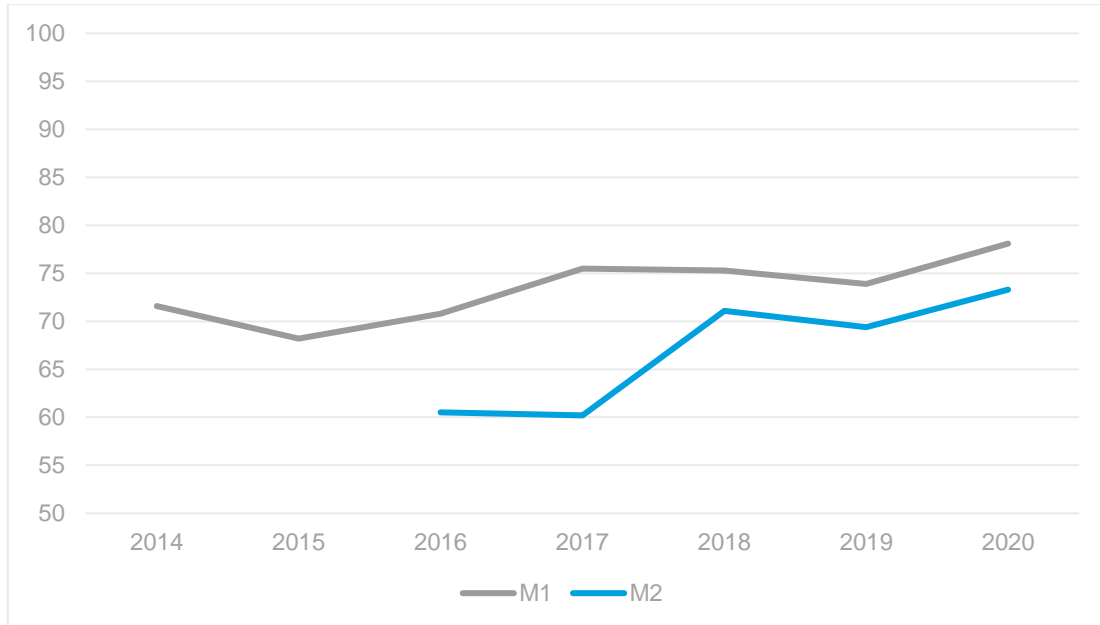
Regarding the two samples M1 and M2 (**Table 2.21**), a slightly higher response rate was achieved for the older sample M1.

Table 2.21: **Response rate by sample (in %)**

	Total	M1	M2
Response rate <sup>1</sup>	76.8	78.1	73.3

<sup>1</sup> RR = percentage of all households with at least one hh and individual interview in the gross sample (gross sample adjusted for households where the last person is deceased or the household moved abroad, is permanently untraceable or dissolved households where the last member moved into another SOEP household).

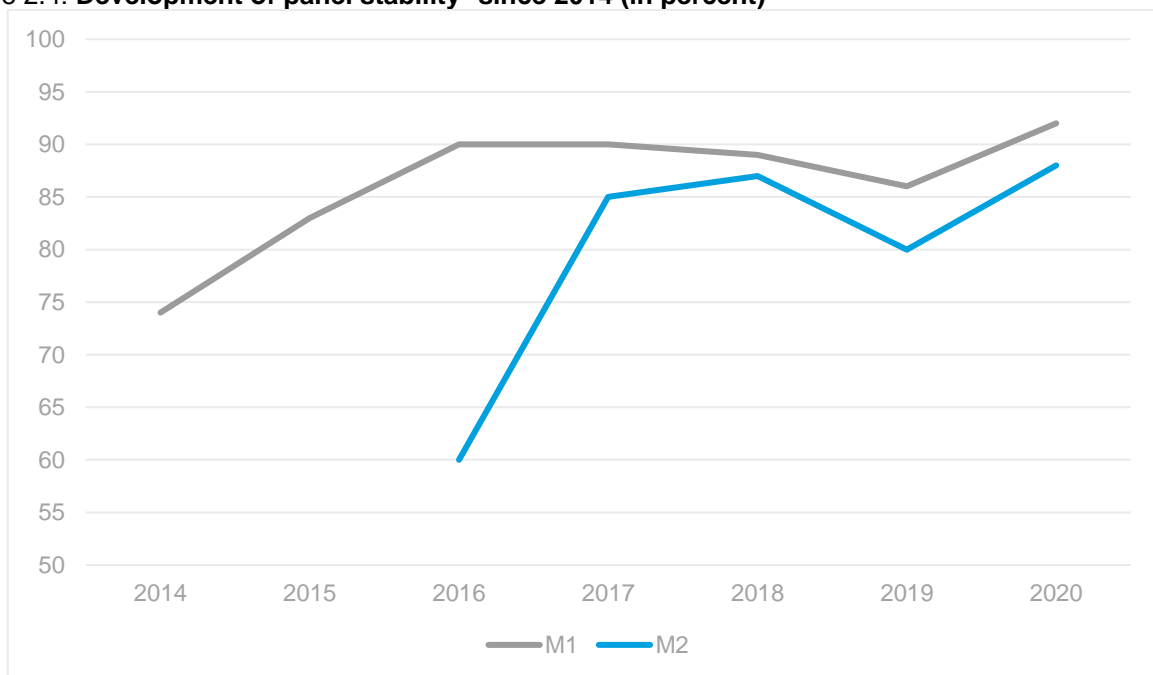
Figure 2.3: Development of response rates<sup>1</sup> since 2014 (in percent)



<sup>1</sup> RR = percentage of all households with at least one hh and individual interview in the gross sample (gross sample adjusted for households where the last person is deceased or the household moved abroad, is permanently untraceable or dissolved households where the last member moved into another SOEP household).

As visualized in **Figure 2.3**, the response rate for sample M2 has been on a lower level than M1 since its inclusion. This year's result of a 73.3 percent response rate for M2 forms a substantial improvement compared to 2017 (60.1), while M1 remained fairly stable over the last three waves.

Figure 2.4: Development of panel stability<sup>1</sup> since 2014 (in percent)



<sup>1</sup> Number of participating households divided by previous wave's net sample.

**Figure 2.4** provides an overview of the development of panel stability for both samples. The panel stability for M1 has been increasing constantly between 2014 and 2016 to then level off at around 90 percent and has been in constant but gradual decline until 2019, then reaching 92.4 percent in 2020. M2's panel stability has been on the rise as well, setting a new record with 87.1 percent in 2018, dropping in 2019 to 80.3 percent, but raise again in 2020 to 88.0 percent.

## 2.6.2 Participation on individual level

Figures for individual-level participation by five different kinds of panel states are presented in **Table 2.22**. Of 4,195 individuals in the gross sample, 54.6 percent were successfully interviewed. Another 28.2 percent dropped out temporarily while 17.2 percent do not want to take part in the survey at all and have expressed this wish either in the current or one of the previous waves. 1,065 individuals dropped out of the survey temporarily in 2019 and were invited one more time to this year's survey. 19.1 percent of them participated. 102 individuals joined existing panel households. 57.8 percent of them provided an interview. Another 54 individuals reached the required age to become a regular panel member this year. 68.5 percent of them participated. Also, we welcomed back 36 persons who had previously stated their permanent refusal to participate back to the panel in 2020 (7.1 percent).

Table 2.22: Participation by panel status (with AAPOR codes)

	Total		Respondents in previous wave		Temp. dropouts in previous wave		Permanent refusals in previous waves		New panel members <sup>1</sup>		Grown into Panel <sup>2</sup>	
	Abs.	In % gross sample	Abs.	In% gross sample	Abs.	In % gross sample	Abs.	In % gross sample	Abs.	In % gross sample	Abs.	In % gross sample
Gross sample <sup>3</sup>	4,195	100.0	2,464	100.0	1,065	100.0	510	100.0	102	100.0	54	100.0
Interview (1.0)	2,292	54.6	1,957	79.4	203	19.1	36	7.1	59	57.8	37	68.5
Non-interview (2.0)	1,903	45.4	507	20.6	861	80.9	474	92.9	43	42.2	17	31.5
Temporary dropout	1,181	28.2	385	15.6	754	70.8	0	0.0	28	27.5	14	25.9
Non-contact (2.20)	12	0.3	9	0.4	3	0.3	0	0.0	0	0.0	0	0.0
Temporary refusal (2.112)	601	14.3	370	15.0	203	19.1	0	0.0	17	16.7	11	20.4
Temporarily physically or mentally unable/incompetent (2.321)	2	0.0	2	0.1	0	0.0	0	0.0	0	0.0	0	0.0
Person could not be traced (temporary) (3.18; 2.4)	7	0.2	4	0.2	3	0.3	0	0.0	0	0.0	0	0.0
Other temp. (2.52)	559	13.3	0	0.0	545	51.2	0	0.0	11	10.8	3	5.6
Final dropout	722	17.2	122	5.0	107	10.1	474	92.9	15	14.7	3	5.6
Permanent refusal (2.111)	680	16.2	101	4.1	88	8.3	474	92.9	14	13.7	3	5.6
Permanently physically or mentally unable/incompetent (2.322)	2	0.0	0	0.0	1	0.1	0	0.0	1	1.0	0	0.0
Deceased (2.31)	4	0.1	1	0.0	3	0.3	0	0.0	0	0.0	0	0.0
Moved abroad (4.2)	19	0.5	12	0.5	7	0.7	0	0.0	0	0.0	0	0.0
Person untraceable (4.4)	17	0.4	8	0.3	9	0.8	0	0.0	0	0.0	0	0.0

<sup>1</sup> New household members who have never been part of the panel.

<sup>2</sup> Former youths who have been part of the panel and take part as official respondents for the first time (with the individual questionnaire).

<sup>3</sup> All household members intended to participate with the individual questionnaire in the current wave, not restricted to members of participating households.

Participation numbers aggregate to a total adjusted response rate of 55.2 percent on the individual level (**Table 2.23**). The rate for previous-wave respondents is noticeably higher at 80.1 percent, while previous-wave dropouts, new panel members and youths produced lower response rates. The permanent refusal conversion rate was 7.1 percent.

Table 2.23: Response rate (in percent)

	Total	Respondents in previous wave	Temp. dropouts in previous wave	Permanent refusals in previous waves	New panel members <sup>2</sup>	Grown into Panel <sup>3</sup>
Response rate <sup>1</sup>	55.2	80.1	19.4	7.1	57.8	68.5

<sup>1</sup> RR = percentage of all participants in the gross sample of individuals born before 2002 (gross sample adjusted for persons who are deceased, moved abroad or are untraceable).

<sup>2</sup> New household members who have never been part of the panel.

<sup>3</sup> Former youths who have been part of the panel and take part as official respondents for the first time (with the individual questionnaire).

A major concern for all SOEP samples is the rate of partial unit non-response (PUNR). A household counts towards PUNR if at least one individual questionnaire is missing. Figures for the survey wave in 2020 are provided in **Table 2.24**. With 29.4 percent, the PUNR was almost equal between M1 and M2 with 29.5 percent.

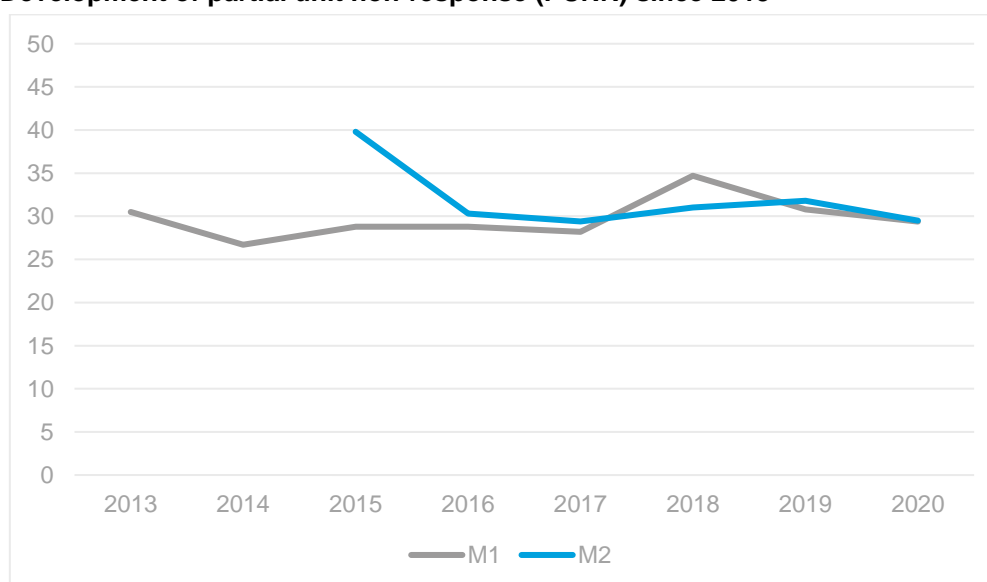
Table 2.24: Partial unit non-response<sup>1</sup> (in percent)

	Total	M1	M2
PUNR <sup>1</sup>	29.4	29.4	29.5

<sup>1</sup> Share of households (number of household members > 1) with at least one missing individual questionnaire.

**Figure 2.5** visualizes how partial unit non-response rates have developed since 2013. Between 2015 and 2017, the rate was fairly constant and below 30 percent for M1. In 2018 experienced a moderate increase. Between 2020 and the previous wave we can observe a drop in PUNR of 1.7 percentage points.

Figure 2.5: Development of partial unit non-response (PUNR) since 2013<sup>1</sup>



<sup>1</sup> Share of households (number of household members > 1) with at least one missing individual questionnaire.

## 2.6.3 Participation by types of questionnaires

**Table 2.25** presents figures on the frequency of each questionnaire. For samples M1 and M2 combined, 1,296 household and 2,292 individual interviews were conducted. Generally, figures are lower for the smaller sample M2. Regarding the youth and child questionnaires, numbers between 45 and 100 are attained. Two sets of numbers are provided for the questionnaire for parents: age 7 or 8. Contrary to the mother and child questionnaires, both parents were asked to answer this questionnaire. The first set of values equals the total amount of questionnaires completed while the second set counts only one questionnaire per child.

Table 2.25: Number of interviews by sample and questionnaire

	Total	M1	M2
Household questionnaire	1,296	952	344
Individual questionnaire	2,292	1,720	572
Life-history module	71	53	18
Youth questionnaire: age 16 or 17	45	34	11
Cognitive competency tests	15	9	6
Youth questionnaire: age 13 or 14	60	53	7
Youth questionnaire: age 11 or 12	48	42	6
Mother and child questionnaire: age 9 or 10	86	66	20
Questionnaire for parents: age 7 or 8 <sup>1</sup>	97	61	36
Questionnaire for parents: age 7 or 8 <sup>2</sup>	60	39	21
Mother and child questionnaire: age 5 or 6	68	39	29
Mother and child questionnaire: age 2 or 3	52	34	18
Mother and child questionnaire: newborn	46	27	19
Questionnaire "Gap"	149	103	46
Questionnaire "Deceased person"	21	19	2

<sup>1</sup> Total number of questionnaires filled out by either mother or father.

<sup>2</sup> Number of children for whom at least one questionnaire was filled out.

Numbers of completed questionnaires for each interviewing mode are provided in **Table 2.26**. The household questionnaire as well as the individual questionnaire with the life-history module were exclusively carried out in CAPI mode, as this is the only mode allowed. Self-interviewing with a paper questionnaire provided by the interviewer was the only available mode for the cognitive competency test and was otherwise relevant



especially for the youth questionnaires for ages 13 or 14 as well as 11 or 12 and for the mother and child questionnaire age 9 or 10 and the parent questionnaire.

Table 2.26: Number of interviews by mode and questionnaire

		Interviewer-Based			Total
		CAPI <sup>3</sup>	PAPI	SELF	
Household questionnaire	abs.	804	6	456	1,266
	In %	63.5	0.5	36.0	100.0
Individual questionnaire	abs.	1,344	22	860	2,226
	In %	60.4	1.0	38.6	100.0
Life-history questionnaire	abs.	43	0	28	71
	In %	60,6	0.0	39,4	100.0
Youth questionnaire: age 16 or 17	abs.	27	0	17	44
	In %	61.4	0.0	38.6	100.0
Cognitive competency tests	abs.	-	-	15	15
	In %	-	-	100.0	100.0
Youth questionnaire: age 13 or 14	abs.	23	2	35	60
	In %	38.3	3.3	58.3	100.0
Youth questionnaire: age 11 or 12	abs.	27	3	17	47
	In %	57.4	6.4	36.2	100.0
Mother and child questionnaire: age 9 or 10	abs.	54	2	28	84
	In %	64.3	2.4	33.3	100.0
Questionnaire for parents: age 7 or 8 <sup>1</sup>	abs.	45	2	45	92
	In %	48.9	2.2	48.9	100.0
Questionnaire for parents: age 7 or 8 <sup>2</sup>	abs.	28	2	28	58
	In %	48.3	3.4	48.3	100.0
Mother and child questionnaire: age 5 or 6	abs.	41	0	0	41
	In %	100.0	0.0	0.0	100.0
Mother and child questionnaire: age 2 or 3	abs.	22	0	0	22
	In %	100.0	0.0	0.0	100.0
Mother and child questionnaire: newborn	abs.	28	1	0	29
	In %	96.6	3.4	0.0	100.0
Questionnaire "Gap"	abs.	119	5	25	149
	In %	79.9	3.4	16.8	100.0
Questionnaire "Deceased person"	abs.	14	0	0	14
	In %	100.0	0.0	0.0	100.0

<sup>1</sup> For missing interviews, in comparison with the previous table, no information for this mode was available.

<sup>2</sup> Total number of questionnaires filled out by either mother or father.

<sup>3</sup> Including CAPI interviews conducted via telephone (see section 2.5.1 for more information)

The response rate for each respective questionnaire is presented in **Table 2.27**. With 2,289 conducted interviews with adults in participating households, the response rate for the individual questionnaire amounts to 69.4 percent. The response rate for the different youth questionnaires and all of the mother and child questionnaires is higher.

Table 2.27: **Response rates by questionnaire**

	Gross sample <sup>1</sup>	Number of interviews	Response rate
Individual questionnaire <sup>2</sup>	2,638	2,289	69.4
Youth questionnaire: age 16 or 17	56	45	80.4
Cognitive competency tests	45	15	33.3
Youth questionnaire: age 13 or 14	68	60	88.2
Youth questionnaire: age 11 or 12	60	48	80.0
Mother and child questionnaire: age 9 or 10	91	86	94.5
Questionnaire for parents: age 7 or 8	65	55	83.3
Mother and child questionnaire: age 5 or 6	71	68	95.8
Mother and child questionnaire: age 2 or 3	56	52	92.9
Mother and child questionnaire: newborn	56	46	82.1
Questionnaire "Gap"	216	149	69.0
Questionnaire "Deceased person"	5	3	60.0

<sup>1</sup> Gross sample= target population in participating households (one household questionnaire and at least one individual questionnaire available), without household members who are deceased or have moved abroad.

<sup>2</sup> Including interviews with first time respondents that included the life-history questionnaire. Four persons filled out an individual interview who were not part of a participating household.

## 2.6.4 Interview length per questionnaire

Figures on the median interview length for each questionnaire are included in **Table 2.28**. A median personal interview in the two samples took 40 minutes, while the household interview took 15 minutes. The youth questionnaires were a little shorter but also had a length of 27 to 33 minutes.

Table 2.28: **Median interview length (minutes) by modes**

	Total		Interviewer-based		
			CAPI	PAPI	SELF
	Median	Percentiles (5%, 95%)	Median	Median	Median
Household questionnaire	15	(7, 50)	13	-	30
Individual questionnaire	40	(20, 90)	37	-	45
Life-history questionnaire	20	(8, 60)	16	-	33
Youth questionnaire: age 16 or 17	33	(19, 60)	34	-	30
Cognitive competency tests	Not recorded				
Youth questionnaire: age 13 or 14	30	(15, 60)	28	-	30
Youth questionnaire: age 11 or 12	27	(11, 58)	25	-	30
Mother and child questionnaire: age 9 or 10	14	(8, 40)	14	-	15
Questionnaire for parents: age 7 or 8	10	(4, 29)	9	-	15
Mother and child questionnaire: age 5 or 6	Not recorded				
Mother and child questionnaire: age 2 or 3	Not recorded				
Mother and child questionnaire: newborn	Not recorded				
Questionnaire "Gap"	2	(1, 15)	1	-	10
Questionnaire "Deceased person"	Not recorded				

## 2.6.5 Consent to record linkage

In 2020, the target group designated for record linkage to employment-history data available at the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB Nuremberg) consisted of 53 participants. They either joined the study for the first time as adults who moved into the households recently, as adults who had lived in the household before but participated in the study for the first time, or as youths that have been part of the panel and take part as official respondents for the first time. (15.1 percent of respondents gave their consent to record linkage (**Table 2.29**).

Table 2.29: **Consent to record linkage IEBS**

	Total	M1	M2
Gross sample <sup>1</sup>	53	37	16
Consent	8	3	5
Consent rate (in %)	15.1	8.1	31.3

<sup>1</sup> First-time participants in the survey.

In addition to the linkage to the IEBS data, all respondents were asked their permission to link their survey data to registration data available at the German statutory pension insurance (Deutsche Rentenversicherung – DRV). 19.4 percent of respondents in the sample agreed to the linkage, as outlined in **Table 2.30**.

Table 2.30: **Consent to record linkage DRV**

	Total	M1	M2
Gross sample <sup>1</sup>	2,214	1,663	551
Consent	429	317	112
Consent rate (in %)	19.4	19.1	20.3

## 2.7 Data preparation

Data preparation processes in samples M1/2 are in line with the processes that are described for samples A-Q in **Section 1.7** of this report.

## 2.8 Delivered data

### Gross Data

Gross data Household  
 Gross data Individuals  
 Interviewer data  
 SOEP Individuals Sample M1+M2

hbru\_M1-2\_2020.dta  
 pbru\_M1-2\_2020.dta  
 intband\_M1-2\_2020.dta  
 P\_M1-2\_2020.SAV

### Net Data

Net data Household checked  
 Net data Household unchecked  
 Net data Individuals checked  
 Net data Individuals unchecked  
 Net data life history checked  
 Net data life history unchecked  
 Net data Youth (age 16 or 17) checked  
 Net data Youth (age 16 or 17) unchecked  
 Net data Cognitive competency test  
 Net data Youth (age 13 or 14)  
 Net data Youth (age 11 or 12)  
 Net data Mother and child E (age 9 or 10)  
 Net data Questionnaire for parents (age 7 or 8)  
 Net data Mother and child C (age 5 or 6)  
 Net data Mother and child B (age 2 or 3)  
 Net data Mother and child A (newborn)  
 „Gap“ data from 2018  
 Net data “Deceased person”

H20\_M1-2.dta  
 H20\_M1-2\_u.dta  
 P20\_M1-2.dta  
 P20\_M1-2\_u.dta  
 L20\_M1-2.dta  
 L20\_M1-2\_u.dta  
 J20\_M1-2.dta  
 J20\_M1-2\_u.dta  
 DJ20\_M1-2.dta  
 FJ20\_M1-2.dta  
 S20\_M1-2.dta  
 ME20\_M1-2.dta  
 EL20\_M1-2.dta  
 MC20\_M1-2.dta  
 MB20\_M1-2.dta  
 MA20\_M1-2.dta  
 LUE19\_20\_M1-2.dta  
 VP20\_M1-2.dta

### Other Data

Professions, sectors, final coding  
 University coding

Berufe\_B Branchen\_Ausb\_  
 2020\_M1M2.dta  
 Hochschul-Abschl\_  
 2020\_M1M2.dta

## 3 Samples M3-5

### 3.1 Introduction

**Table 3.1** gives a short overview of the main characteristics of the 2020 wave of samples M3-5. 2,408 households from samples M3-5 participated between August 2020 and February 2021. This results in an adjusted response rate of 69.5 percent. Participation of households that were respondents in the previous wave was at 76.7 percent. In all households, 3,166 adults (without first-time respondents) and 220 youths gave interviews. For an additional 1,272 children of various age groups, data are available from interviews with a parent. 13 different questionnaires were fielded in CAPI mode. No other modes were used in M3-5 in 2020. Partial unit non-response (PUNR<sup>7</sup>) was at 66.3 percent.

Table 3.1: **Summary fieldwork**

Fieldwork period	August 2020 – February 2021
Mode (main questionnaires)	CAPI
Gross sample (hh)	3,330
Net sample (hh)	2,408
Response rate (adjusted; hh) <sup>1</sup>	Overall: 76.7 Prev. wave respondents: 82.5 Prev. wave dropouts: 47.5 New households: 63.0
Number of questionnaires	Adults: 4 Youths: 3 Children: 5
Net sample (individuals)	Adults: 3,166 Youths: 220 Children: 1,272
Questionnaire length (median, in minutes)	Household: 20 Adult <sup>2</sup> : 45
Partial unit non-response (PUNR)	66.3

<sup>1</sup> RR = percentage of all households with at least one household and individual interview in the gross sample (gross sample adjusted for households where the last person is deceased or the household moved abroad, is permanently untraceable, or dissolved households where the last member moved into another SOEP household).

<sup>2</sup> Without first time respondents.

<sup>7</sup> PUNR: share of households (number of household members > 1) with at least one missing individual questionnaire.

### 3.2 Development samples M3-5

The refugee samples were integrated into SOEP-Core to represent the rising number of refugees who arrived in Germany since 2013 (**Figure 3.1**). To implement an innovative sampling procedure mapping recent migration and integration dynamics, the SOEP partnered with the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB Nuremberg) and the Research Centre of the Federal Office for Migration and Refugees (Forschungszentrum des Bundesamt für Migration und Flüchtlinge, BAMF-FZ) in 2016. M3 is the acronym for the first boost sample of 1,673 households of adult refugees who entered Germany from January 1, 2013 to January 31, 2016 and applied for asylum in Germany (**Figure 3.2**). M4 is the acronym for the second refugee boost sample with 1,601 households. It consists of two tranches. The first one is a household boost of the M3 sample. For the second tranche, underage children of refugee families were sampled, but only the adults in the respective households were invited to participate. M5 is the acronym for the third boost sample of 1,519 refugee households that was established in 2017. The population covers adult refugees who have applied for asylum in Germany since January 1, 2013. For all three samples, the Central Register of Foreign Nationals (Ausländerzentralregister, AZR) was utilized as a sampling frame.<sup>8</sup> In 2020, the fourth wave of sample M5 and the fifth waves of samples M3 and M4 were fielded.

Figure 3.1: Overview of refugee samples

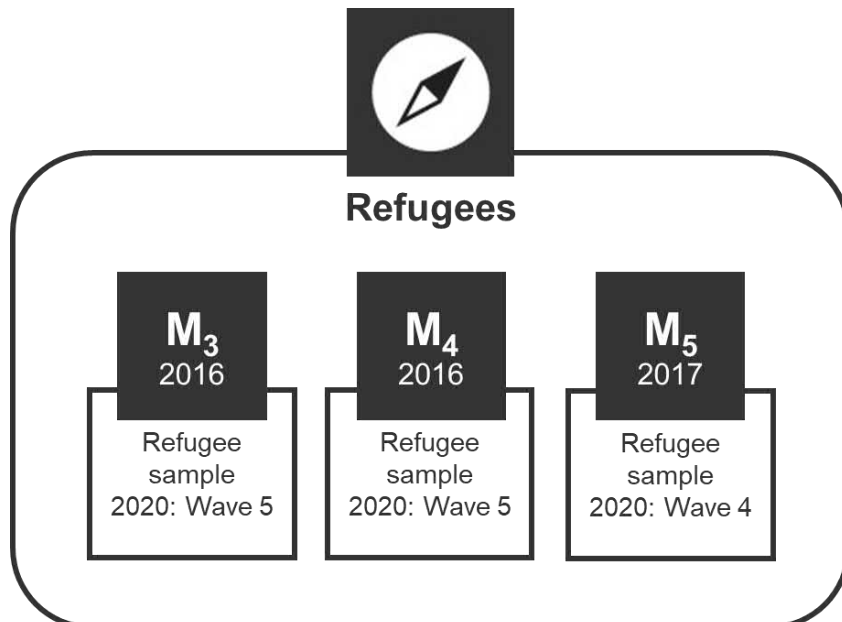
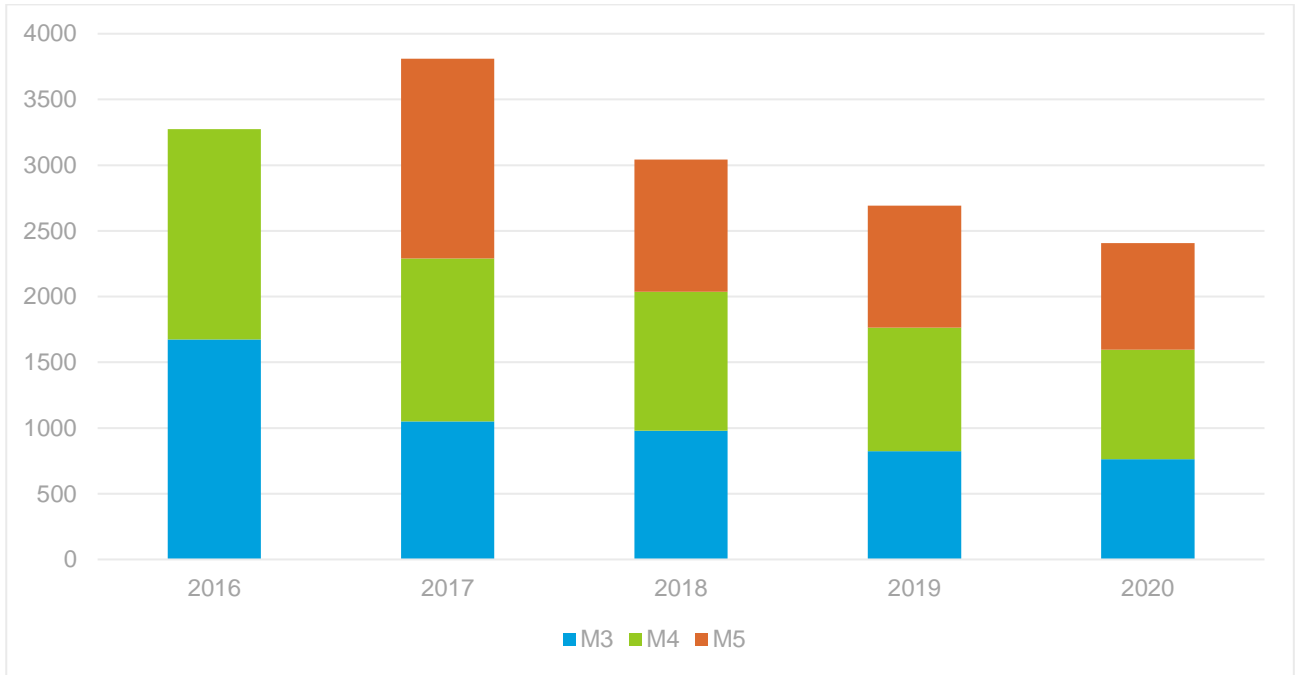


Figure 3.2: Development of net samples M3-5 since 2016 (number of households)

<sup>8</sup> The sampling design of the refugee samples M3 and M4 is described in: SOEP Wave Report 2016; the sampling design for M5 in: SOEP Wave Report 2017.





## 3.3 Structure of the gross sample

The final gross sample of the previous year serves as a starting point for the following year's survey wave. Of 4,133 households in samples M3-5 in 2019, 64.9 percent were coded as respondents without refusal for the next wave and are thus expected to participate in 2020 (**Table 3.2**). A total of 21.8 percent of all households in the samples dropped out permanently in 2019. 13.0 percent of households dropped out temporarily. For all samples, no households were excluded for dropping out temporarily in two consecutive waves.

Table 3.2: Final gross sample 2019 by type of household

Household level	Total		M3		M4		M5	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample</b>	4,133	100.0	1,272	100.0	1,359	100.0	1,502	100.0
Participating households without refusal for next wave	2,681	64.9	817	64.2	937	68.9	927	61.7
Participating households with refusal for next wave	12	0.3	6	0.5	4	0.3	2	0.1
Temporary dropouts	537	13.0	180	14.2	140	10.3	217	14.4
Final dropouts	903	21.8	269	21.1	278	20.5	356	23.7
2-year rule	0	0.0	0	0.0	0	0.0	0	0.0

**Table 3.3** presents the gross sample of 2020 for three different types of household. Of 3,330 households in the gross sample, 80.9 percent were respondents in the previous wave. 537 households in M3-5 which dropped out temporarily in 2019 are invited to take part in 2020 again. Additionally, 100 new split-off households were identified and included in the samples.

Table 3.3: Administered gross sample 2020 by type of household

Household level	Total		M3		M4		M5	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample</b>	3,330	100.0	1,037	100.0	1,107	100.0	1,186	100.0
Participating households in previous wave	2,693	80.9	823	79.4	941	85.0	929	78.3
Temporary dropout in previous wave	537	16.1	180	17.4	140	12.6	217	18.3
New households (split-off hhs)	100	3.0	34	3.3	26	2.3	40	3.4

Key characteristics of the households in samples M3-5 are presented in **Table 3.4** and **Table 3.5**. Most frequently, households consist of five or more household members (36.5 percent). However, there is also a notable number of households over all three samples which consist of only one household member (29.2 percent). 33.1 percent are positioned in or around larger cities.

Table 3.4: Household characteristics by sub-samples I

Household level	Total		M3		M4		M5	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample</b>	3,230	100.0	1,003	100.0	1,081	100.0	1,146	100.0
<b>HH size<sup>1</sup></b>								
1	943	29.2	344	34.3	189	17.5	410	35.8
2	286	8.9	98	9.8	81	7.5	107	9.3
3	333	10.3	117	11.7	90	8.3	126	11.0
4	488	15.1	149	14.9	166	15.4	173	15.1
5+	1,180	36.5	295	29.4	555	51.3	330	28.8
<b>State</b>								
Schleswig-Holstein	149	4.6	50	5.0	66	6.1	33	2.9
Hamburg	80	2.5	43	4.3	18	1.7	19	1.7
Lower Saxony	392	12.1	99	9.9	120	11.1	173	15.1
Bremen	46	1.4	32	3.2	13	1.2	1	0.1
North Rhine-Westphalia	775	24.0	182	18.1	214	19.8	379	33.1
Hesse	310	9.6	79	7.9	86	8.0	145	12.7
Rhineland Palatinate	137	4.2	32	3.2	37	3.4	68	5.9
Baden-Wuerttemberg	372	11.5	106	10.6	160	14.8	106	9.2
Bavaria	396	12.3	132	13.2	141	13.0	123	10.7
Saarland	65	2.0	43	4.3	22	2.0	0	0.0
Berlin	116	3.6	61	6.1	33	3.1	22	1.9
Brandenburg	107	3.3	47	4.7	47	4.3	13	1.1
Mecklenburg Western Pomerania	39	1.2	16	1.6	16	1.5	7	0.6
Saxony	104	3.2	18	1.8	61	5.6	25	2.2
Saxony-Anhalt	62	1.9	20	2.0	35	3.2	7	0.6
Thuringia	80	2.5	43	4.3	12	1.1	25	2.2

<sup>1</sup> Status as of previous wave; new households are consequently missing (all household members including children).

Table 3.5: Household characteristics by sub-samples II

Household level	Total		M3		M4		M5	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample<sup>1</sup></b>	3,230	100.0	1,003	100.0	1,081	100.0	1,146	100.0
<b>BIK type<sup>2</sup></b>								
0	809	25.0	268	26.7	239	22.1	302	26.4
1	262	8.1	31	3.1	122	11.3	109	9.5
2	784	24.3	277	27.6	247	22.8	260	22.7
3	424	13.1	117	11.7	139	12.9	168	14.7
4	94	2.9	18	1.8	43	4.0	33	2.9
5	243	7.5	107	10.7	60	5.6	76	6.6
6	354	11.0	108	10.8	134	12.4	112	9.8
7	197	6.1	63	6.3	73	6.8	61	5.3
8	43	1.3	9	0.9	18	1.7	16	1.4
9	20	0.6	5	0.5	6	0.6	9	0.8
<b>Community size<sup>3</sup></b>								
1	48	1.5	14	1.4	13	1.2	21	1.8
2	147	4.6	30	3.0	63	5.8	54	4.7
3	656	20.3	180	17.9	247	22.8	229	20.0
4	730	22.6	197	19.6	280	25.9	253	22.1
5	444	13.7	181	18.0	116	10.7	147	12.8
6	703	21.8	209	20.8	227	21.0	267	23.3
7	502	15.5	192	19.1	135	12.5	175	15.3

<sup>1</sup> Status as reported at the end of wave 2019. New households in 2020 are consequently missing.

<sup>2</sup> BIK-type: 0 (more than 500,000 inhabitants/center), 1 (more than 500,000 inh./periphery), 2 (100,000 to 499,999 inh./center), 3 (100,000 to 499,999 inh./periphery), 4 (50,000 to 99,999 inh./center), 5 (50,000 to 99,999 inh./periphery), 6 (20,000 to 49,999 inh.), 7 (5,000 to 19,999 inh.), 8 (2,000 to 4,999 inh.), 9 (fewer than 2,000 inh.).

<sup>3</sup> Community size: 1 (fewer than 2000 inhabitants), 2 (2,000 to 5,000 inh.), 3 (5,000 to 20,000 inh.), 4 (20,000 to 50,000 inh.), 5 (50,000 to 100,000 inh.), 6 (100,000 to 500,000 inh.), 7 (more than 500,000 inh.).

## 3.4 Questionnaires and fieldwork material

### 3.4.1 Questionnaires

One special feature of samples M3-5 is the use of three different versions of individual questionnaires (**Table 3.6**). For M3-5, a special SOEP individual questionnaire was developed that covers issues specific to refugees. First time respondents received a longer version of the refugee questionnaire that includes the life-history questionnaire. With the help of a short screener, other members of the household could be identified as refugees or non-refugees. Non-refugees received the same questionnaire as adult respondents from samples A-Q and M1/2, including the life-history module. The life-history module was an additional part at the end of the questionnaire only for first time respondents. Also, there are youth and child questionnaires with some specific refugee issues for the same age groups as in the samples A-Q and M1/2. All questionnaires are solely available in CAPI mode and provided in six different language versions (see **Section 3.5.3**).

Table 3.6: **Questionnaires and modes**

	CAPI
Household questionnaire	✓
Individual questionnaire for refugees	✓
Individual questionnaire + life-history questionnaire for refugees	✓
Individual questionnaire (+ life-history module) for non-refugees	✓
Youth questionnaire: age 16 or 17	✓
Youth questionnaire: age 13 or 14	✓
Youth questionnaire: age 11 or 12	✓
Mother and child questionnaire: age 9 or 10	✓
Questionnaire for parents: age 7 or 8	✓
Mother and child questionnaire: age 5 or 6	✓
Mother and child questionnaire: age 2 or 3	✓
Mother and child questionnaire: newborn	✓

At the household level, in addition to the standard household questionnaire, a mother-child questionnaire was used, merging the questionnaires previously used for children of different age groups.

One notable feature of this year's questionnaire was the escape-route map. It had already been used the years before in samples M3-5. It was integrated in the questionnaire for first-time respondents aimed at refugees. The escape-route map is a tool to reconstruct a refugee's route from their home country to their arrival in

Germany. A world map was presented to the respondents. By clicking on the screen, the respondents could select their home country and they could mark all stops along their route. They were urged to select not only countries but also to mark all important cities and border crossing points as well.

As with every previous subsample of the migration population in the SOEP, questionnaire content was based on the SOEP-Core questionnaires. However, there were several deviations from the SOEP-Core standard to reflect the special characteristics of the target group, including several additional questions on migration and integration.

## 3.4.2 Fieldwork material

In addition to the questionnaires, a whole range of fieldwork materials such as letters, leaflets or documents for the interviewers were designed, printed, and sent to households and interviewers. **Table 3.7** provides an overview of the different material types that were prepared in samples M3-5. Because the mode was restricted to CAPI in these samples, the number of different versions of materials was notably smaller than in samples A-H. However, many materials were provided in seven different languages (German, English, English, Arabic, Farsi, Pashtu, Urdu, and Kurmanji).

### **Advance letter**

About two weeks before the start of the fieldwork period, the households received an advance letter in which the interviewer's visit was announced. It was always sent in German language as well as in one other of the six other available languages in accordance with the language version chosen for the interview in 2019. The letter included a link to the SOEP website that provided additional information in the different languages.

### **Leaflet**

Every household in samples M3-5 received an eight-page leaflet with reports and published results specifically from the refugee sample. The leaflets in German and the second language were sent with the advance letter.

### **Declaration on data protection**

Every household got access to a two-page declaration on data protection detailing the organizations that are responsible for processing all respondent data along with a description of data handling and data recipients. The declaration on data protection in German and the second language was sent with the advance letter.

### **Consent to record linkage form**

Two different consent to record linkage were integrated. First, to all new respondents born in 2002 or later a consent to record linkage form was presented that allows an individual linkage between the respondent's data and employment history data available at the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB Nuremberg). Respondents who have been asked in one of the previous waves but declined or did not understand the issue were asked for their consent a second time.

In addition to the linkage to the IAB data, respondents who had told us in this wave or in any previous wave that they had participated in an integration course of the German Federal Office for Migration and Refugees were asked their permission to link their survey data to registration data available in the "Integration Business File" (Integrationsgeschäftsdatei, or InGe) of the BAMF.

Table 3.7: **Fieldwork material**

CAPI	
Advance letter	<ul style="list-style-type: none"> <li>• In 7 languages (German, English, Arabic, Farsi, Pashtu, Urdu, Kurmanji)</li> </ul>
Leaflet	<ul style="list-style-type: none"> <li>• In 7 languages</li> </ul>
Declaration on data protection	<ul style="list-style-type: none"> <li>• In 7 languages</li> </ul>
Consent to record linkage form <sup>1</sup>	<ul style="list-style-type: none"> <li>• In 7 languages</li> </ul>
Address form and household grid <sup>2</sup>	<ul style="list-style-type: none"> <li>• Electronic form (Mein Kantar)</li> </ul>
Project instruction book	<ul style="list-style-type: none"> <li>• Version for M3-5</li> </ul>
Other interviewer material	<ul style="list-style-type: none"> <li>• Project description</li> <li>• Contact card (in 7 languages)</li> <li>• HH information card</li> </ul>
Additional interviewer material for M3-5 to use in the households	All in 7 languages: <ul style="list-style-type: none"> <li>• How is a survey carried out?</li> <li>• FAQ</li> <li>• Aid to fill out the household grid</li> <li>• Postcard for movers</li> </ul> In German with subtitles in Arabic, Farsi or English: <ul style="list-style-type: none"> <li>• Motivational film</li> </ul>
Additional interviewer material for M3-5 to use when dealing with shared accommodations	<ul style="list-style-type: none"> <li>• Letter from the Federal Office for Migration and Refugees (BAMF)</li> <li>• Leaflet describing the survey</li> </ul>

<sup>1</sup> Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB Nuremberg): Link to employment history data / German Federal Office for Migration and Refugees (BAMF): Link to Integration Business File.

<sup>2</sup> Including the so-called "B3 form" used to process address changes.

### Address form and household grid

The address form provided an overview of the household composition as it was last known to Kantar. The interviewers had to document when and with whom the survey was conducted or why a sample member did not participate in the current year. They were also asked to note every single contact attempt made. For samples M3-5, interviewers did this in electronically in the "Mein Kantar" software.

Moreover, the interviewers were asked to carefully document any moves of households and household members or changes in the household composition. An additional form had to be completed for every person that had left the household since the last survey because all of these remain potential respondents and were tracked even if the new address was unknown.

### Project instruction book

In addition to a shorter project description, interviewers in samples M3-5 also received a much more detailed instruction manual that is about 70 pages long. This manual contained information on special features of the current wave, specific processing instructions, and questionnaires as well as background information on the project.

## **Other interviewer material**

Moreover, the interviewers received contact cards for households that could not be reached at home and household information cards with information on individual numbers, names, years of birth, types of questionnaires, incentives, survey modes, and notes.

## **Additional interviewer material to use in M3-5 households**

Because the language barriers in households of samples M3-5 were often higher than in other samples, the interviewers received a few additional laminated sheets in all seven languages to help explain how a survey is carried out to address frequently asked questions and to help fill out the household grid. For movers, the interviewers also left a postcard with the households that asks them to send their new address to Kantar.

In 2020 Kantar provided a short motivational film to the interviewers that features one of our interviewers who is a refugee from Syria himself. The film was integrated into the household questionnaire and could be played in German with subtitles in German, Arabic, Farsi, or English as required.

## **Additional interviewer material to use when dealing with shared accommodations**

In samples M3-5, interviewers might have to deal with employees and security personnel at shared refugee accommodations. For these cases, the Federal Office for Migration and Refugees (Bundesamt für Migration und Flüchtlinge, BAMF) sent the interviewers a letter and a leaflet describing the survey to be handed over to shared accommodations.

## 3.5 Conducting the survey

### 3.5.1 Survey mode

All questionnaires used in samples M3-5 were solely available in CAPI mode. Hence, no interviews in other modes such as SELF interviewing, which was possible for youth and child questionnaires in samples M1/2, existed in samples M3-5 in 2020. However, because of the coronavirus pandemic, interviewers and respondents were given the option to conduct the interviews via telephone (CAPI via telephone or CAPI-TEL). This option was taken in 13.5 percent of all cases on the household level and 12.6 percent of cases on the individual level (refugees only).

### 3.5.2 Fieldwork timing

The fieldwork progress for each month is depicted in **Table 3.8**. Fieldwork started in August and lasted until February of 2021. September and December were the most productive months with more than 20 percent of the net samples being processed in each of these months. Fieldwork was extended for several reasons until February 2021. Many addresses were no longer accurate for fieldwork and required further research. It was also more difficult to find times when respondents could meet with interviewers than in the first wave, and many appointments had to be rescheduled.

Table 3.8: **Monthly fieldwork progress**

Household level	Gross Sample		Net Sample	
	Abs.	In %	Abs.	In %
August <sup>1</sup>	200	6.0	154	6.4
September	594	17.8	506	21.0
October	554	16.6	455	18.9
November	361	10.8	285	11.8
December	712	21.4	519	21.6
January	471	14.1	297	12.3
February	438	13.2	192	8.0

<sup>1</sup> Including households that refused to take part in the survey prior to start of fieldwork.



## 3.5.3 Translations

Language problems during the interviewing process form a potential major challenge for surveys with populations that recently entered the country as refugees. Although some of the interviewers conducting in M3-5 speak Arabic, Farsi, or Pashtu, it is generally not feasible to match interviewers with special language skills with respondents in such a large, nationwide survey. As implemented successfully in the first wave of samples M3/4 in 2016, a bilingual CAPI program was used for all three refugee samples in 2020. Questions in German and a second language were presented on the screen side by side as shown in **Figure 3.3**. The language was selected at the beginning of the interview. Respondents were given the additional choice to have no translations on screen (**Table 3.9**).

Figure 3.3: Screenshot bilingual CAPI program – language selection

The screenshot shows a language selection screen with the following text and options:

Welche Sprachversion soll verwendet werden?  
 Which language combination should be used for the interview?  
 ما هو الزوج اللغوي الذي تريد استخدامه لإجراء الحوار؟  
 مايل هستيد كه مصاحبه به چه زباني انجام بگيرد؟  
 کومه ژبه بايد د مرکي لياره وکارول شي؟  
 الترويو کسے لئے کون سی زبانوں کا امتزاج استعمال کیا جانا چاہئے؟

Kijan ziman bikar bê?

- 1. Deutsch / Englisch  
1 German/English
- 2. Deutsch / Arabisch  
2 ألماني / عربي
- 3. Deutsch / Farsi  
3 ألماني / فارسي
- 4. Deutsch / Paschtu  
4 ألماني/پشتو
- 5. Deutsch / Urdu  
5 جرمن / اردو
- 6. Deutsch / Kurmandschi  
6 Almanî / Kurmancî
- 7. Deutsch / ohne Übersetzung

Table 3.9: Available language versions

	German / English	German / Arabic	German / Farsi	German / Pashtu	German / Urdu	German / Kurmanji	German only
Household questionnaire	✓	✓	✓	✓	✓	✓	✓
Individual questionnaire for refugees	✓	✓	✓	✓	✓	✓	✓
Individual + life-history questionnaire for refugees	✓	✓	✓	✓	✓	✓	✓
Youth questionnaires	✓	✓	✓	✓	✓	✓	✓
Mother and child questionnaire/ questionnaire for parents	✓	✓	✓	✓	✓	✓	✓

**Table 3.10** presents the utilization of each language version for the individual questionnaires. A translated individual questionnaire was used for 3,181 interviews. With 84.9 percent, the German / Arabic version was used most frequently, followed by the German / Farsi version with 8.2 percent. The questionnaires in Pashtu, Farsi, and Kurmanji were selected far less frequently. The option to use no translation was used for 3.2 percent of the interviews with individual questionnaire for refugees.

In addition to the questionnaires and the fieldwork materials that were available in the different languages, the interviewers could call a so-called “interpreter hotline” during the process of contacting the households. In these cases, other interviewers that are fluent in either Arabic or Farsi helped the interviewer, e.g., to explain the study background, answer questions and set up an appointment for the interview. However, this service could only be used to convince respondents to participate in the study. The interview itself needed to be conducted with the bilingual questionnaires.

**Table 3.10: Utilization of a certain language version – individual questionnaire for refugees<sup>1</sup>**

	Total		Individual questionnaire		Individual questionnaire + life-history questionnaire	
	Abs.	In %	Abs.	In %	Abs.	In %
German / English	68	2.1	68	2.2	-	-
German / Arabic	2,700	84.9	2,618	84.7	82	90.1
German / Farsi	261	8.2	255	8.3	6	6.6
German / Pashtu	13	0.4	13	0.4	-	-
German / Urdu	20	0.6	19	0.6	1	1.1
German / Kurmanji	18	0.6	18	0.6	-	-
Without any translation	101	3.2	99	3.2	2	2.2
<b>Total</b>	<b>3,181</b>	<b>100.0</b>	<b>3,090</b>	<b>100.0</b>	<b>91</b>	<b>100.0</b>

<sup>1</sup> Individual questionnaire and individual questionnaire + life-history questionnaire for new respondents.

For the 76 non-refugees who received the individual questionnaire from samples A-Q and M1/2, translated paper versions in five different languages including English, Russian, Turkish, Romanian, and Polish were available (**Table 3.11**). Assistance with language problems was needed in four cases, which amounts to 5.3 percent (**Table 3.12**). The person in question used a polish paper questionnaire for assistance with the interview.

Table 3.11: Translated paper questionnaires for the individual questionnaires for non-refugees

	English	Russian	Turkish	Romanian	Polish
Individual questionnaire for non-refugees	✓	✓	✓	✓	✓
Life-history questionnaire for non-refugees	✓	✓	✓	✓	✓

Table 3.12: Language problems and usage of translated paper questionnaires<sup>1</sup> for non-refugees

	Total	In % net sample
Net sample (individual questionnaire) <sup>2</sup>	76	100.0
No language problems occurred/no need for assistance with language problems	72	94.7
Assistance with language problems needed	4	5.3
Of that number:		
German-speaking person in the same household	3	3.9
German-speaking person from outside the household	0	0.0
Professional interpreter	1	1.3
Translated paper questionnaire	0	0.0

<sup>1</sup> Individual questionnaire + life-history questionnaire for non-refugees.

<sup>2</sup> Including all individual questionnaires even if the household in which they are conducted is classified as a non-participating household.

## 3.5.4 Panel maintenance and incentives

In the first waves of samples M3-5, households did not receive any cash incentives or vouchers as many households still lived in shared accommodations and might experience problems when presented with cash by the interviewers. At the beginning of fieldwork in 2018, interviewers were given a choice between handing over a notebook with a pen as a small gift or bringing a small gift of their choice of less than 5 euros in value. After a couple of weeks of fieldwork, some interviewers reported that participants increasingly asked for cash incentives because some had heard that other surveys provided them. Consequently, in 2020 interviewers were given the option to incentivize households that answer one household and one individual questionnaire with 15 euros either in cash or in the form of a voucher or gifts. The interviewer was free to choose the form of the incentive. The distribution of these choices across samples M3-5 is detailed in **Table 3.13**. In 93.1 percent of cases the cash incentive was paid, only 2.0 percent of households received vouchers. Another 4.9 percent of households rejected any incentive in 2020.

Table 3.13: Incentives samples M3-5

Household level	Total		M3		M4		M5	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
Cash-incentive	2,242	93.1	719	94.1	752	90.4	771	94.8
Voucher	49	2.0	4	0.5	36	4.3	9	1.1
No Incentive	118	4.9	41	5.4	44	5.5	33	4.9

## 3.5.5 Movers and tracing

Looking at the gross sample, 25.7 percent of all households in samples M3-5 were identified as movers or new households and therefore, their new addresses needed to be traced (**Table 3.14**). It is not surprising that this share is much higher than in samples A-Q (9.1 percent). The tracing of these households was successful in 79.0 percent of all cases. Most new addresses were gained by post-address research and interviewer efforts.

Table 3.14: Movers and sources of new addresses of administered sample 2020

Household level	Total	
	Abs.	In %
<b>Gross sample</b>	3,330	100.0
Movers and new households	857	25.7
<b>Success tracing</b>		
Tracing successful	677	79.0
Tracing not successful	180	21.0
<b>Source</b>		
Interviewer	262	32.3
Postal service	63	7.8
Local registration offices	1	0.1
Participant	9	1.1
Client	0	0.0
Post-Address Research	476	58.7

### 3.5.6 Interviewer characteristics, training & monitoring

**Table 3.15** presents key characteristics of the interviewers working in the samples M3-5. A total of 53 interviewers were active over all three samples and 83.0 percent of them were male. A high share of the interviewers was between 21 and 40 years old. Thus, the age structure is younger compared to the interviewing staff in the SOEP-Core samples. Most interviewers processed between 20 and 99 households.

Table 3.15: Interviewer characteristics

Interviewer level	Total		M3		M4		M5	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Number of Interviewers</b>	53	100.0	48	100.0	49	100.0	49	100.0
Gender								
Male	44	83.0	40	83.3	41	83.7	41	83.7
Female	9	17.0	8	16.7	8	16.3	8	16.3
Age								
21-39 years	30	56.6	29	60.4	27	55.1	29	59.2
40-59	19	35.8	16	33.3	18	36.7	16	32.7
60-79	4	7.5	3	6.3	4	8.2	4	8.2
Number of households (gross)								
Fewer than 5	5	9.4	10	20.8	10	20.4	5	10.2
5 – 19	8	15.1	14	29.2	12	24.5	23	46.9
20 – 99	30	56.6	24	50.0	27	55.1	21	42.9
More than 99	10	18.9	0	0.0	0	0.0	0	0.0

Because fieldwork in samples M3-5 was carried out by a special group of interviewers of whom many can speak Arabic, the staff was not only notably different in age compared to the interviewers in samples A-Q. They were also much less experienced as interviewers, as shown in **Table 3.16**. 92.5 percent had less than 5 years of experience working for Kantar compared to only 17.3 percent in samples A-Q.

Table 3.16: Interviewer experience

Interviewer Level	Total		M3		M4		M5	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
<b>Number of Interviewers</b>	53	100.0	48	100.0	49	100.0	49	100.0
<b>Experience with Kantar</b>								
0-4 years of experience with Kantar	49	92.5	45	93.8	45	91.8	45	91.8
5-9 years	2	3.8	1	2.1	2	4.1	2	4.1
10-19 years	1	1.9	1	2.1	1	2.0	1	2.0
20-29 years	1	1.9	1	2.1	1	2.0	1	2.0
<b>Experience with SOEP</b>								
0-4 years	51	96.2	47	97.9	47	95.9	47	95.9
5-9 years	2	3.8	1	2.1	2	4.1	2	4.1

All interviewers in samples M3-5 were trained face-to-face by members of the project team at Kantar a couple of days prior to the start of fieldwork. Topics of the one-day event held in August 2020 were:

- Welcome and overview of the survey
- General processing rules and special features of SOEP samples M3-5
- Field documents
- Overview of the questionnaires
- Using the sample management system “Mein Kantar”
- Using the CAPI survey software “Compass 32”
- Central organizational aspects of field organization
- Tutorial

Kantar places high priority on interviewer monitoring and has put an ISO-certified process in place that is audited regularly. Kantar adheres to the German Business Association of Market and Social Research Institutes (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V., ADM) standards for internal regulation and monitoring of all systems and procedures. This means that a minimum of 10 percent of Kantar’s annual interviews are checked, and every interviewer is monitored at least once a year.

In 2018, we expanded these existing quality control measures for interviewer monitoring in the SOEP projects. The basic interviewer-monitoring concept was expanded to a project-based control system for all SOEP samples: A number of participating households is contacted shortly after the interview by letter or phone asking them to confirm their participation in a regularly conducted interview. In case of inconsistencies and/or irregularities, we attempted to gain clarification through direct contact with respondents, primarily by telephone.

**Table 3.17** shows details for the re-contacting process in samples M3-5. Of 2,408 households in the net sample, 80.8 percent were re-contacted after the interviews had been conducted in order to assure data quality and identify non-standard behavior by interviewers. In terms of survey mode, phone interviews were the standard approach to ensure a relatively high response rate. Only households for which no telephone number was available received a short paper questionnaire.

Overall, we received feedback from 1,389 households. This results in a response rate of 71.4 percent overall. Using the results from the re-contacting process as well as analyses of paradata and interview data, we identified no interviewer who had not adhered to our standards in conducting interviews in the 2020 fieldwork period of M3-5.

Table 3.17: Interviewer monitoring

	Total		
	Abs.	In %	In %
By households			
Households in net sample	2,408	100.0	-
Re-contacted households	1,945	80.8	100.0
Households with feedback	1,389	57.7	71.4
By interviewers			
Interviewers in net sample	53	100.0	-
Interviewers with re-contacted households	50	94.3	-
Interviewers with non-standard behavior	0	0.0	-

## 3.6 Fieldwork results

### 3.6.1 Participation on household level

**Table 3.18** provides a detailed overview on participation figures for three different types of households. In total, 2,408 of 3,330 households in the gross sample were interviewed. 79.1 percent of respondents participated in the previous wave as well while only 42.1 of pre-wave dropouts did so in 2020. Also 51 new households were successfully interviewed which amounts to an overall response rate of 51.0 percent.

Table 3.18: Participation by type of household (with AAPOR codes)

	Total		Respondents in previous wave		Dropouts in previous wave		New households	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
Gross sample	3,330	100.0	2,693	100.0	537	100.0	100	100.0
<b>Interview (1.0)</b>	<b>2,408</b>	<b>72.3</b>	<b>2,131</b>	<b>79.1</b>	<b>226</b>	<b>42.1</b>	<b>51</b>	<b>51.0</b>
Complete (1.1)	1,301	39.1	1,138	42.3	129	24.0	34	34.0
Partial (1.2)	1,107	33.2	993	36.9	97	18.1	17	17.0
<b>Non-interview (2.0)</b>	<b>922</b>	<b>27.7</b>	<b>562</b>	<b>20.9</b>	<b>311</b>	<b>57.9</b>	<b>49</b>	<b>49.0</b>
<b>Temporary dropout</b>	<b>371</b>	<b>11.1</b>	<b>224</b>	<b>8.3</b>	<b>125</b>	<b>23.3</b>	<b>22</b>	<b>22.0</b>
Non-contact (2.20)	279	8.4	169	6.3	100	18.6	10	10.0
Temporarily physically or mentally unable (2.321)	4	0.1	2	0.1	2	0.4	0	0.0
Temporary refusal (2.351; 2.353; 2.354)	59	1.8	43	1.6	16	3.0	0	0.0
Household could not be traced (temporary) (3.18; 2.4)	27	0.8	10	0.4	6	1.1	11	11.0
Not attempted or worked (3.11)	2	0.1	0	0.0	1	0.2	1	1.0
<b>Final dropout</b>	<b>551</b>	<b>16.5</b>	<b>338</b>	<b>12.6</b>	<b>186</b>	<b>34.6</b>	<b>27</b>	<b>27.0</b>
Permanent refusal (2.111)	347	10.4	217	8.1	122	22.7	8	8.0
Deceased (2.31)	3	0.1	2	0.1	1	0.2	0	0.0
Permanently physically or mentally unable (2.322)	12	0.4	9	0.3	3	0.6	0	0.0
Language problem (2.331)	2	0.1	2	0.1	0	0.0	0	0.0
Moved abroad (4.2)	7	0.2	3	0.1	4	0.7	0	0.0
Household dissolved (4.3)	1	0.0	1	0.0	0	0.0	0	0.0
Household untraceable (4.4)	179	5.4	104	3.9	56	10.4	19	19.0



11.1 percent of the gross sample dropped out temporarily with the household being unreachable as most frequent reason. Another 16.5 percent of the gross sample dropped out of the survey permanently.

Table 3.19: Participation by sample (with AAPOR codes)

	Total		M3		M4		M5	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
Gross sample	3,330	100.0	1,037	100.0	1,107	100.0	1,186	100.0
<b>Interview (1.0)</b>	<b>2,408</b>	<b>72.3</b>	<b>764</b>	<b>73.7</b>	<b>832</b>	<b>75.2</b>	<b>812</b>	<b>68.5</b>
Complete (1.1)	1,301	39.1	413	39.8	422	38.1	466	39.3
Partial (1.2)	1,107	33.2	351	33.8	410	37.0	346	29.2
<b>Non-interview (2.0)</b>	<b>922</b>	<b>27.7</b>	<b>273</b>	<b>26.3</b>	<b>275</b>	<b>24.8</b>	<b>374</b>	<b>31.5</b>
<b>Temporary dropout</b>	<b>371</b>	<b>11.1</b>	<b>104</b>	<b>10.0</b>	<b>106</b>	<b>9.6</b>	<b>161</b>	<b>13.6</b>
Non-contact (2.20)	279	8.4	83	8.0	72	6.5	124	10.5
Temporary refusal (2.351; 2.353; 2.354)	59	1.8	10	1.0	23	2.1	26	2.2
Temporarily physically or mentally unable (2.321)	4	0.1	2	0.2	2	0.2	0	0.0
Household could not be traced (temporary) (3.18; 2.4)	27	0.8	9	0.9	9	0.8	9	0.8
Not attempted or worked (3.11)	2	0.1	0	0.0	0	0.0	2	0.2
<b>Final dropout</b>	<b>551</b>	<b>16.5</b>	<b>169</b>	<b>16.3</b>	<b>169</b>	<b>15.3</b>	<b>213</b>	<b>18.0</b>
Permanent refusal (2.111)	347	10.4	114	11.0	113	10.2	120	10.1
Permanently physically or mentally unable (2.322)	12	0.4	4	0.4	4	0.4	4	0.3
Language problem (2.331)	2	0.1	0	0.0	0	0.0	2	0.2
Deceased (2.31)	3	0.1	0	0.0	1	0.1	2	0.2
Moved abroad (4.2)	7	0.2	2	0.2	1	0.1	4	0.3
Household dissolved (4.3)	1	0.0	0	0.0	1	0.1	0	0.0
Household untraceable (4.4)	179	5.4	49	4.7	49	4.4	81	6.8

Numbers of participation differentiated for the three samples M3-5 are presented in **Table 3.19**. The interviewing rates vary slightly between samples while sample M5 has the lowest interviewing rate. The rates of temporary and permanent dropouts are also slightly different across all three samples.

**Table 3.20** presents the overall adjusted response rate as well as response rates for the different types of households mentioned earlier. Adjusted response means that gross sample adjusted for households where the last person is deceased or the household moved abroad, is permanently untraceable, or dissolved

households where the last member moved into another SOEP household. The overall adjusted response rate of all households in samples M3-5 amounts to 76.7 percent compared to 69.5 percent in 2019. With 82.5 percent it is distinctly higher for households which also responded in last year's survey compared to 47.5 percent for the group of pre-wave dropouts. New households come in with a response rate of 63.0 percent. Response rates for all three samples M3-5 are provided in **Table 3.21**. M4 and M5 have the highest and lowest adjusted response rates with 78.9 percent and 73.9 percent, respectively.

Table 3.20: **Response rate by type of household (in percent)**

	Total	Respondents in previous wave	Drop-outs in previous wave	New households
Response rate <sup>1</sup>	76.7	82.5	47.5	63.0

<sup>1</sup> RR = percentage of all households with at least one household and individual interview and households in the gross sample (gross sample adjusted for households where the last person is deceased or the household moved abroad, is permanently untraceable, or dissolved households where the last member moved into another SOEP household).

Table 3.21: **Response rate by sample (in percent)**

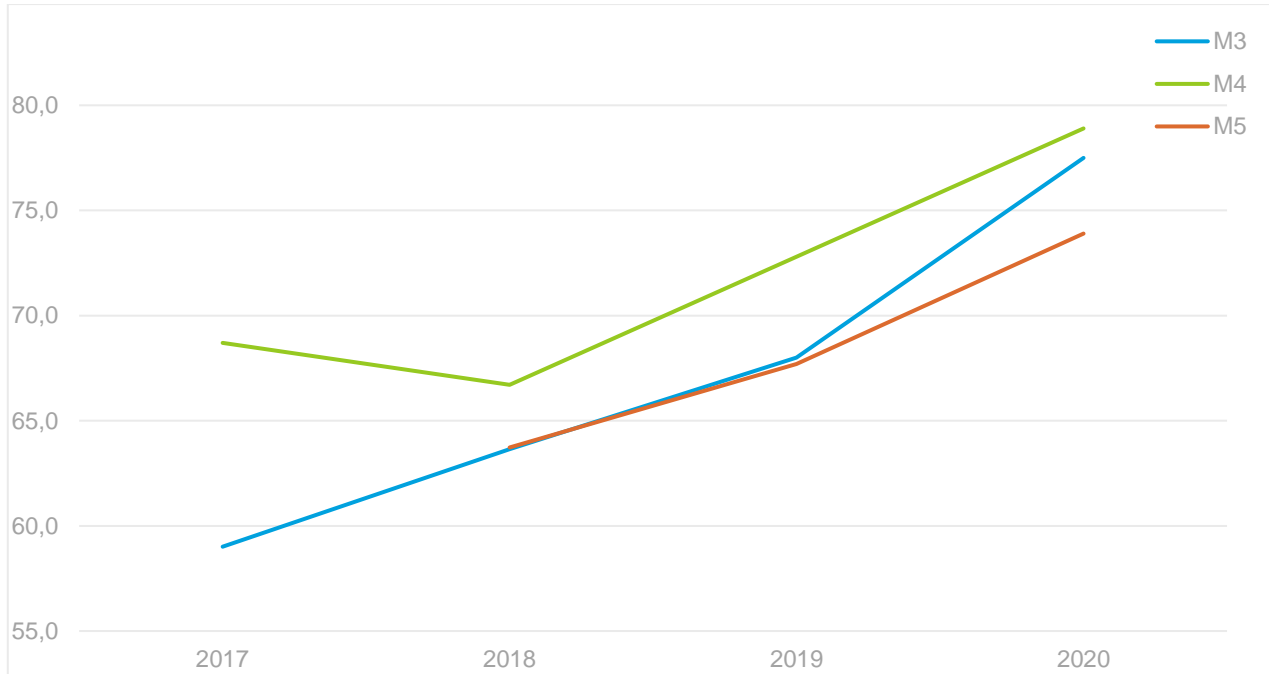
	Total	M3	M4	M5
Response rate <sup>1</sup>	76.7	77.5	78.9	73.9

<sup>1</sup> RR = percentage of all households with at least one household and individual interview and households in the gross sample (gross sample adjusted for households where the last person is deceased or the household moved abroad, is permanently untraceable, or dissolved households where the last member moved into another SOEP household).

The development of response rates since 2017 is presented in **Figure 3.4**. While M3 and M5 both show a continuous increase since their respective first waves, the response rate for M4 decreased by approximately 2 percentage points in 2018. Since then, the response rate of M4 increased to 78.9 percent in 2020. Over all waves, M4 shows a higher response rate than M3 and M5.

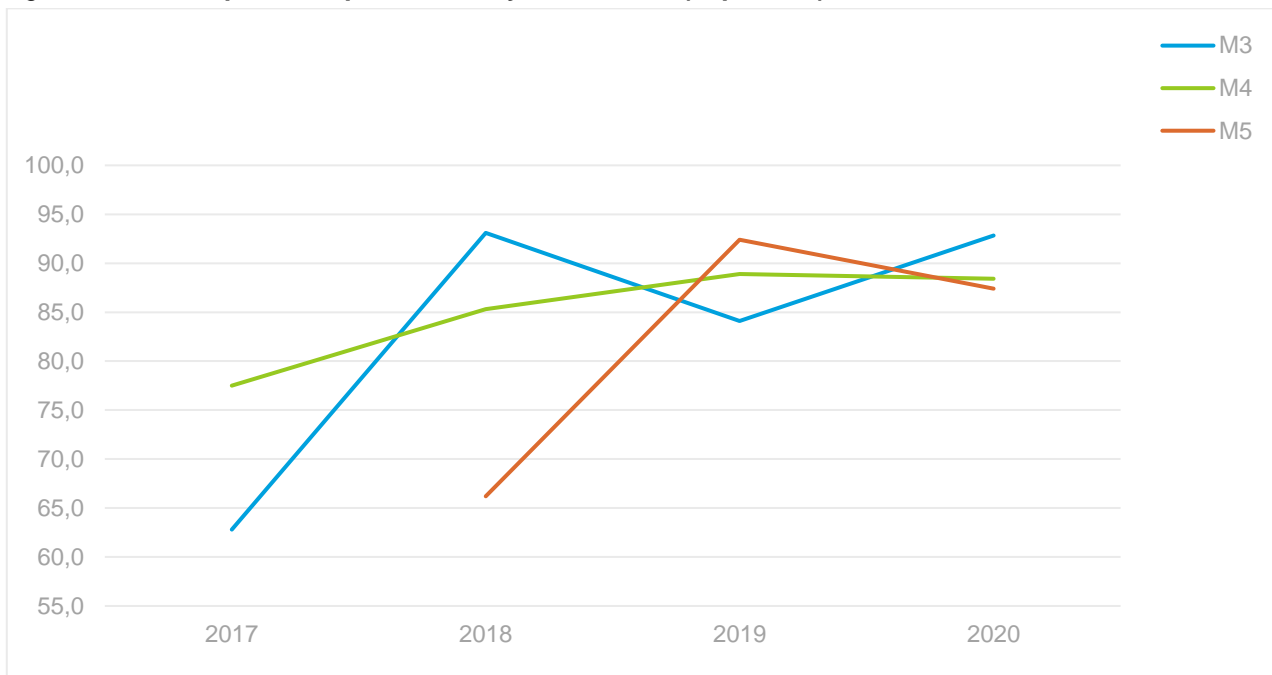
In terms of panel stability, M4 was the only sample with a continuously increasing panel stability between waves until 2019 but stagnating with 88.4 percent in 2020 (**Figure 3.5**). While M3 scored 14.5 percent lower on stability than M4 in 2017, the sample reached high levels with 93.0 percent in 2018 and 92.8 percent in 2020, surpassing M4. M5 declined from 92.4 percent in 2019 to 87.4 percent in 2020. In order to meaningfully assess panel stability rates over the years, the various subsamples should be processed for at least five consecutive waves due to specific gross sample processes in the first couple of waves. After this time period, the panel stability rates of samples are usually consolidated and therefore comparable.

Figure 3.4: Development of response rates<sup>1</sup> since 2017 (in percent)



<sup>1</sup> RR = percentage of all households with at least one household and individual interview and households in the gross sample (gross sample adjusted for households where the last person is deceased or the household moved abroad, is permanently untraceable, or dissolved households where the last member moved into another SOEP household).

Figure 3.5: Development of panel stability<sup>1</sup> since 2017 (in percent)



<sup>1</sup> Number of participating households divided by previous wave's net sample.

## 3.6.2 Participation on individual level

**Table 3.22** presents participation numbers by panel status on individual level. Of 6,995 individuals in the gross sample, 46.6 percent were successfully interviewed, while 53.4 percent either dropped out temporarily (18.7 percent) or permanently (34.8 percent). The participation rate of individuals who also participated in the previous wave was 66.9 percent. Of all pre-wave dropouts, only 32.4 percent were willing to participate again in the present year. The rate for new panel members was 16.0 percent and 43.1 percent for youths. In 2020, 96 people of those who had permanently refused participation in the past were successfully re-integrated into the panel resulting in a permanent refusal conversion rate of 8.1 percent.

Table 3.22: Participation by panel status (with AAPOR codes)

	Total		Respondents in previous wave		Temp. dropouts in previous wave		Permanent refusals in previous waves		New panel members <sup>1</sup>		Grown into Panel <sup>2</sup>	
	Abs.	In % gross sample	Abs.	In% gross sample	Abs.	In % gross sample	Abs.	In % gross sample	Abs.	In % gross sample	Abs.	In % gross sample
Gross sample <sup>3</sup>	6,995	100.0	3,845	100.0	1,561	100.0	1,186	100.0	331	100.0	72	100.0
Interview (1.0)	3,257	46.6	2,571	66.9	506	32.4	96	8.1	53	16.0	31	43.1
Non-interview (2.0)	3,738	53.4	1,274	33.1	1,055	67.6	1,090	91.9	278	84.0	41	56.9
Temporary dropout	1,305	18.7	632	16.4	505	32.4	0	0.0	146	44.1	22	30.6
Temporary refusal (2.112)	1,155	16.5	561	14.6	451	28.9	0	0.0	122	36.9	21	29.2
Non-contact (2.21)	16	0.2	5	0.1	8	0.5	0	0.0	3	.9	0	.0
Temporarily physically or mentally unable/incompetent (2.321)	56	0.8	37	1.0	17	1.1	0	0.0	2	0.6	0	0.0
Language problem (2.331)	9	0.1	4	0.1	4	0.3	0	0.0	1	0.3	0	0.0
Other temp. (2.52)	35	0.5	10	0.3	6	0.4	0	0.0	18	5.4	1	1.4
Person could not be traced (temporary) (3.18; 2.4)	34	0.5	15	0.4	19	1.2	0	0.0	0	0.0	0	0.0
Final dropout	2,433	34.8	642	16.7	550	35.2	1,090	91.9	132	39.9	19	26.4
Permanent refusal (2.111)	2,146	30.7	491	12.8	433	27.7	1,090	91.9	124	37.5	8	11.1
Deceased (2.31)	5	0.1	4	0.1	1	0.1	0	0.0	0	0.0	0	0.0
Permanently physically or mentally unable/incompetent (2.322)	4	0.1	1	0.0	3	0.2	0	0.0	0	0.0	0	0.0
Moved abroad (4.2)	16	0.2	9	0.2	7	0.4	0	0.0	0	0.0	0	0.0
Person untraceable (4.4)	262	3.7	137	3.6	106	6.8	0	0.0	8	2.4	11	15.3

<sup>1</sup> New household members who have never been part of the panel.

<sup>2</sup> Youths who have been part of the panel and take part as official respondents for the first time (with the youth questionnaire age 16/17).

<sup>3</sup> All household members intended to participate with the adult questionnaire in the current wave, not restricted to members of participating households.

Response rates by panel status are provided in **Table 3.23**. In difference to participation rates as presented in **Table 3.22**, response rates are calculated excluding deceased individuals, individuals who have moved abroad, or who are untraceable. The overall response rate is 48.5 percent, which is a 2.2 percentage point decrease compared to the previous wave. This decrease can be attributed to a significantly lower permanent refusal conversion rate and a significantly lower response rate among new panel members in the wave of 2020 compared to the previous wave.

Table 3.23: Response rate (in percent)

	Total	Respondents in previous wave	Temp. dropouts in previous wave	Permanent refusals in previous waves	New panel members <sup>1</sup>	Grown into Panel <sup>2</sup>
Response rate <sup>3</sup>	48.5	69.6	35.0	8.1	16.4	50.8

<sup>1</sup> New household members who have never been part of the panel.

<sup>2</sup> Former youths who have been part of the panel and take part as official respondents for the first time (with the individual questionnaire).

<sup>3</sup> RR = percentage of all participants in the gross sample of individuals born before 2002 (gross sample adjusted for persons who are deceased, moved abroad or are untraceable).

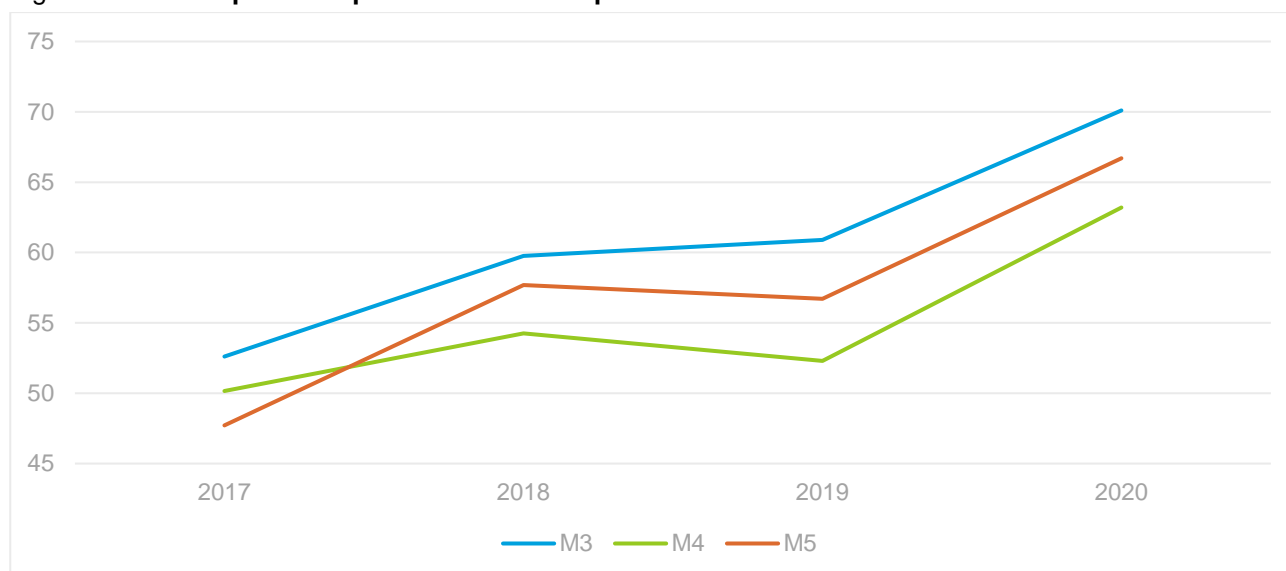
One major concern for all SOEP samples are the growing partial unit non-response (PUNR) rates, which are exceptionally high for the refugee samples, at a total of 66.3 percent in this year's wave (**Table 3.24**). According to our interviewers' reports, respondents are increasingly busy with activities such as job search, participation in language and integration courses, and appointments with various agencies and authorities. The increasing number of activities these individuals are involved in makes it difficult for interviewers to complete interviews with multiple adult household members. **Figure 3.6** visualizes how partial unit non-response rates increased between 2017 and 2018, stabilized in 2019 but again increased in 2020 over all samples.

Table 3.24: Partial unit non-response (in percent)

	Total	M3	M4	M5
PUNR <sup>1</sup>	66.3	70.1	63.2	66.7

<sup>1</sup> Share of households (number of household members > 1) with at least one missing individual questionnaire.

Figure 3.6: Development of partial unit non-response since 2017<sup>1</sup>



<sup>1</sup> Share of households (number of household members > 1) with at least one missing individual questionnaire.

## 3.6.3 Participation by types of questionnaires

As presented in **Table 3.25**, all in all 2,408 household questionnaires and 3,090 individual questionnaires were completed. Three different versions of individual questionnaires are to be distinguished as mentioned earlier. The individual questionnaire for refugees was used most frequently over all three samples. Furthermore, information for 220 youths in three age groups between 11 and 17 years and for 1,272 children in five age groups from newborns to children of 10 years of age is available.

Table 3.25: Number of interviews by sample and questionnaire

	Total	M3	M4	M5
Household questionnaire	2,408	764	832	812
Individual questionnaire for refugees	3,090	933	1,140	1,017
Individual questionnaire + life-history questionnaire for refugees	91	30	23	38
Individual questionnaire (+ life-history module) for non-refugees <sup>1</sup>	76 (51)	16 (9)	48 (33)	12 (9)
Youth questionnaire: age 16 or 17	62	62	62	62
Youth questionnaire: age 13 or 14	81	81	81	81
Youth questionnaire: age 11 or 12	77	77	77	77
Mother and child questionnaire: age 9 or 10	266	76	123	67
Questionnaire for parents: age 7 or 8 <sup>2</sup>	228	55	105	68
Mother and child questionnaire: age 5 or 6	229	49	102	78
Mother and child questionnaire: age 2 or 3	243	65	80	98
Mother and child questionnaire: newborn	171	100	105	101

<sup>1</sup> The life-history module was an additional part at the end of the questionnaire for non-refugees and only for first time respondents. Number of interviews including the life-history module are in brackets.

<sup>2</sup> In samples M3-5, the questionnaire for parents is answered by one parent but not both as in other samples such as in samples A-Q.

Unlike the response rates on the individual level that were shown in the previous section, response rates for each questionnaire are measured by looking at members of participating households only (**Table 3.26**). Combining all versions of the individual questionnaires, a response rate of 66.3 percent was obtained. Response rates for the youth questionnaires were 32.5 percent on average. All mother and child questionnaires generated very high response rates each above 90 percent.

Table 3.26: **Response rates by questionnaire**

	Gross sample <sup>1</sup>	Number of interviews	Response rate
Individual questionnaire <sup>2</sup>	4,912	3,257	66.3
Youth questionnaire: age 16 or 17	178	62	34.8
Youth questionnaire: age 13 or 14	250	81	32.4
Youth questionnaire: age 11 or 12	248	77	31.0
Mother and child questionnaire: age 9 or 10	281	266	94.7
Questionnaire for parents: age 7 or 8	245	228	93.1
Mother and child questionnaire: age 5 or 6	248	229	92.3
Mother and child questionnaire: age 2 or 3	260	243	93.5
Mother and child questionnaire: newborn <sup>3</sup>	199	171	85.9

<sup>1</sup> Gross sample = target population in participating households without household members who are deceased or have moved abroad. For the child-related questionnaires, the reference value is the number of children in the respective age group living in participating households. Therefore, the response rate for these questionnaires indicates the number of children for whom a questionnaire has been completed by one parent (in most cases by the mother).

<sup>2</sup> Including interviews with first-time respondents.

<sup>3</sup> Gross sample is adjusted for children of the respective age whose parents had already answered the questionnaire in 2019.

## 3.6.4 Interview length per questionnaire

**Table 3.27** lists the median interview lengths for each questionnaire. The median interview length for refugees who had taken part in one of the previous waves was 45 minutes for the individual questionnaire. This meant that the interview was moderately longer than the interviews in other SOEP samples (e.g., 40 minutes in A-Q), adding further to issues of low response rates and high PUNR.



Table 3.27: Median interview length (in minutes)

	Median	Percentiles (5%, 95%)
Household questionnaire	20	(10, 40)
Individual questionnaire for refugees	45	(20, 70)
Individual questionnaire + life-history questionnaire for refugees	50	(24, 77)
Individual questionnaire for non-refugees	38	(20, 73)
Life-history module for non-refugees	18	(3, 44)
Youth questionnaire: age 16 or 17	35	(15, 69)
Youth questionnaire: age 13 or 14	40	(20, 60)
Youth questionnaire: age 11 or 12	30	(12, 55)
Mother and child questionnaire: age 9 or 10	Not directly recorded <sup>1</sup>	
Questionnaire for parents: age 7 or 8		
Mother and child questionnaire: age 5 or 6		
Mother and child questionnaire: age 2 or 3		
Mother and child questionnaire: newborn		

<sup>1</sup> It existed only one mother and child questionnaire with loops for all children in the household. Therefore, the individual interview length could be only calculated with the delivered timestamps.

## 3.6.5 Consent to record linkage

In all SOEP migration and refugee samples, one aim is to link respondents' survey data with data from the Integrated Employment Biographies Sample (Stichprobe der Integrierten Erwerbsbiografien, IEBS) for as many participants as possible. All first-time refugee respondents in 2020 as well as participants who refused once before were asked to give their written consent to this record linkage. As presented in **Table 3.28**, 164 of 211 individuals (77.7 percent) gave their consent.

Table 3.28: Consent to record linkage IEBS

	Total	M3	M4	M5
Gross sample <sup>1</sup>	211	53	92	66
Consent	164	40	76	48
Consent rate (in %)	77.7	75.5	82.6	72.7

<sup>1</sup> In 2020, the gross sample for record linkage consisted of first-time respondents and participants from previous waves who either once refused or did not understand the issue.

In addition to the linkage to the IEBS data, respondents who had told us in this wave or in any previous wave that they had participated in an integration course of the German Federal Office for Migration and Refugees were asked their permission to link their survey data to registration data available in the “Integration Business File” (Integrationsgeschäftsdatei, or InGe) of the BAMF. 80.2 percent of respondents in the gross sample agreed to the linkage, as outlined in **Table 3.29**.

Table 3.29: **Consent to record linkage InGe**

	Total	M3	M4	M5
Gross sample <sup>1</sup>	283	69	91	123
Consent	227	58	76	93
Consent rate (in %)	80.2	84.1	83.5	75.6

## 3.7 Data preparation

Data preparation processes in samples M3-5 are in line with the processes described for samples A-Q in **Section 1.7** of this report.

## 3.8 Delivered data

### Gross Data

Gross data Household  
Gross data Individuals  
Interviewer data  
SOEP Individuals Sample M3-M5

Hbru\_M345\_2020.dta  
Pbru\_M345\_2020.dta  
Intband\_M345\_2020.dta  
P\_M345\_2020.dta

### Net Data

Net data Household checked  
Net data Household unchecked  
Net data Individuals Refugees checked  
Net data Individuals Refugees unchecked  
Net data Individual + Life history Refugees checked  
Net data Individual + Life history Refugees unchecked  
Net data Individuals Non-refugees checked  
Net data Individuals Non-refugees unchecked  
Net data Life history Non-refugees checked  
Net data Life history Non-refugees unchecked  
Net data Youth (all age groups)  
Net data Children (all ages)  
Consent to record linkage  
Aggregated net data Refugee Route Module

H20\_M345.sav, H20\_M345.dta  
H20\_M345\_u.sav, H20\_M345\_u.dta  
P20\_M345.sav, P20\_M345.dta  
P20\_M345\_u.sav, P20\_M345\_u.dta  
PB20\_M345.sav, PB20\_M345.dta  
PB20\_M345\_u.dta  
PDM20\_M345.dta  
PDM20\_M345\_u.dta  
LDM20\_M345.dta  
LDM20\_M345\_u.dta  
J20\_M345.sav, J20\_M345.dta  
MK20\_M345.sav, MK20\_M345.dta  
EV20\_DIW\_M345.dta  
PB20\_Fluchtrouten\_M345.dta

### Other Data

University coding  
Professions, sectors, final coding  
  
Additional Codes for Individual + Life history Refugee

Hochschul-Abschl\_2020\_M345.dta  
BerufeBranchenAusbildung\_2020\_M345.dta  
Nationen.xlsx  
Zusatzcodes\_2020.xlsx

## 4 Sample M6

### 4.1 Introduction

**Table 4.1** gives an overview of the main characteristics of the first wave of the refugee boost sample M6 of 2020. The gross sample for foeldwirk consisted of 3,000 households, 1,141 of whom participated in the survey between August and February. This results in an adjusted response rate of 44.3 percent. As is usually the case for boost and refreshment samples in the SOEP, only questionnaires in CAPI mode were fielded. No other modes were used in M6, although a small percentage of interviews was conducted via telephone in the CAPI environment. Partial unit non-response (PUNR) was at 91.1 percent.

Table 4.1: **Summary fieldwork**

Fieldwork period	August - February
Mode (main questionnaires)	CAPI
Gross sample (hh)	3,000
Net sample (hh)	1,141
Response rate (adjusted; hh) <sup>1</sup>	Overall: 44.3
Number of questionnaires	Adults: 3 Youths: 0 Children: 0
Net sample (individuals)	Adults: 1,216 Youths: 0 Children: 0
Questionnaire length (median, in minutes)	Household: 17 Adult: 60
Partial unit non-response (PUNR)	91.1

<sup>1</sup> RR = percentage of all households with at least one hh and individual interview in the gross sample (gross sample adjusted for households where the anchor respondent is deceased, moved abroad, is permanently untraceable, not attempted or worked).

## 4.2 Background Sample M6

M6 is the acronym for the fourth top-up sample for households that represents refugees. The population of M6 covers two groups: firstly, adult refugees who arrived in Germany between January 1, 2013 and December 2016 (“Refreshment”) and secondly adult refugees who came to Germany between January 1, 2017 and June 2019 (“Enlargement”) with a strongly disproportionate oversampling of refugees from East- and West-Africa.

As with the previous refugee samples M3, M4, and M5 the Central Register of Foreign Nationals (Ausländerzentralregister – AZR) was utilized as sampling frame. The sampling frame of the Central Register of Foreign Nationals provides only basic information about foreigners in Germany, including: nationality, name, date of birth, and a registration number linked to the local recordkeeping authority. Thus, the SOEP group at DIW Berlin and the BAMF-FZ were responsible for the implementation of the sample design and gross sample-selection. The local recordkeeping authorities were the primary sampling units (PSU) in accordance with strata based on the information of the Central Register of Foreign Nationals. By this means 10,207 anchor-respondents were sampled in 159 PSUs by the SOEP group at DIW Berlin and the BAMF-FZ. Of this gross sample 63.5% belonged to the refreshment and 36.5% to the enlargement. The sampled cases were not distributed evenly across the PSUs, the number of cases per PSU varied from two to 1,149 cases. In a second step Kantar sampled 3,000 anchor-respondents for fieldwork. The sample drawn by Kantar was stratified proportionally according to nationality, sex, region of origin (East Africa, West Africa, remaining regions), federal state, BIK-type, and sub-sample (refreshment, enlargement). Therefore, of the 3,000 anchor-respondents in the fieldwork-sample 1,906 belonged to the refreshment and the other 1,094 to the enlargement sample. Due to the low numbers of gross cases in some of the PSUs and to fulfil the strata specifications named above a total of 12 PSUs could not be sampled into the fieldwork sample.

## 4.3 Structure of the Gross Sample

As described in the previous chapter the gross sample for fieldwork consisted of 3,000 households and was proportionally sampled from the gross sample provided by the SOEP group at DIW Berlin and the BAMF-FZ (e.g. regarding federal state and BIK-type). As presented in **Table 4.2**, there is a concentration of households in North Rhine-Westphalia and Berlin. Very few households were located in Brandenburg, Mecklenburg Western Pomerania, Saxony-Anhalt and Thuringia. Besides, the majority of households is positioned in the center of large cities (**Table 4.3**) with either more than 500,000 inhabitants or with 100,000 to 499,999 inhabitants.

Table 4.2: Household characteristics I

Household Level	Total	
	Abs.	In %
<b>Gross Sample<sup>1</sup></b>	<b>3,000</b>	<b>100.0</b>
<b>State</b>		
Schleswig-Holstein	95	3.2
Hamburg	193	6.5
Lower Saxony	266	8.9
Bremen	104	3.5
North Rhine-Westphalia	1,011	33.9
Hesse	198	6.6
Rhineland Palatinate	85	2.9
Baden-Wuerttemberg	166	5.6
Bavaria	211	7.1
Saarland	93	3.1
Berlin	334	11.2
Brandenburg	21	0.7
Mecklenburg Western Pomerania	21	0.7
Saxony	118	4.0
Saxony-Anhalt	32	1.1
Thuringia	34	1.1

<sup>1</sup> For 18 households that moved abroad this information is unavailable

Table 4.3: Household characteristics II

Household Level	Total	
	Abs.	In %
<b>Gross Sample<sup>1</sup></b>	<b>2,982</b>	<b>100.0</b>
<b>BIK-Type<sup>2</sup></b>		
0	1,548	51.9
1	176	5.9
2	541	18.1
3	258	8.7
4	35	1.2
5	113	3.8
6	164	5.5
7	126	4.2
8	17	0.6
9	4	0.1
<b>Community size<sup>3</sup></b>		
1	16	0.5
2	52	1.7
3	408	13.7
4	443	14.9
5	235	7.9
6	595	20.0
7	1,233	41.3

<sup>1</sup> For 18 households that moved abroad this information is unavailable

<sup>2</sup> BIK type: 0 (more than 500,000 inhabitants/center) 1 (more than 500,000 inh/periphery), 2 (100,000 to 499,999 inh./center), 3 (100,000 to 499,999 inh./periphery), 4 (50,000 to 99,999 inh.(center), 5 (50,000 to 99,999 inh./periphery), 6 (20,000 to 49,999 inh.), 7 (5,000 to 19,999 inh.), 8 (2,000 to 4,999 inh.), 9 (fewer than 2,000 inh.)

<sup>3</sup> Community size: 1 (fewer than 2000 inhabitants), 2 (2,000 to 5,000 inh.), 3 (5,000 to 20,000 inh.), 4 (20,000 to 50,000 inh.), 5 (50,000 to 100,000 inh.), 6 (100,000 to. 500,000 inh.), 7 (more than 500,000 inh.).

## 4.4 Questionnaires and fieldwork material

### 4.4.1 Questionnaires

In the first wave of M6 three questionnaires were fielded: the individual questionnaire for first time respondents (including additional biographical questions) for all adult household members, which was administered in separate versions for refugees and for Germans or migrants respectively, and the household questionnaire for the anchor respondent. Like for the other refugee samples M3-5, a special SOEP individual and life-history questionnaire was developed that includes issues specific to refugees. The version for Germans and migrants was identical to the individual and life-history questionnaire in samples A-Q and M1/2. As is the usual approach for boost samples, no youth or child questionnaires were fielded in sample M6. All questionnaires were solely available in CAPI mode and provided in seven different language versions (see **Section 4.5.3**).

Table 4.4: **Questionnaires and modes sample M6**

	CAPI
Household questionnaire	✓
Individual questionnaire + life history questionnaire module I	✓
Individual questionnaire + life history questionnaire module II	✓

In addition to the questionnaires for respondents, interviewers were asked to complete a short questionnaire about the area the household is located in, the so-called “residential environment questionnaire”.



## 4.4.2 Fieldwork material

In addition to the questionnaires, a whole range of fieldwork materials such as letters, leaflets, and documents for the interviewers were designed, printed, and sent to households and interviewers. **Table 4.5** provides an overview of the different material types that were prepared in sample M6. Because the mode was restricted to CAPI, the number of different versions of materials was notably smaller than in samples A-Q. However, many materials were provided in eight different languages (German, English, French, Arabic, Farsi, Pashtu, Urdu, and Kurmanji).

### **Advance letter**

About two weeks before the start of the fieldwork period, households received an advance letter in which the interviewer's visit was announced. It was always sent in German language as well as in one other of the seven other available languages based on the likelihood of the individual speaking a certain language. The letter included a link to the SOEP website that provided additional information in the different languages.

### **Brochure**

Every household received an eight-page brochure with general information on the survey. The brochures in German and the second language were sent with the advance letter.

### **Declaration on data protection**

Every household got access to a two-page declaration on data protection detailing the organizations that are responsible for processing all respondent data along with a description of data handling and data recipients. The declaration on data protection in German and the second language was sent with the advance letter.

### **Consent to record linkage form**

To all respondents born in or before 2002, a consent to record linkage form was presented that allows an individual linkage between the respondent's data and employment history data available at the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB Nuremberg).

### **Hygiene flyer**

In the face of challenges to face-to-face-interviewing brought on by the Coronavirus pandemic, Kantar implemented strict guidelines for interviewers and households alike to ensure the safety of everyone involved. These guidelines were contained in a flyer that the interviewers brought into the household with them.

Table 4.5: Fieldwork material sample M6

CAPI	
Advance letter	<ul style="list-style-type: none"> <li>• In 8 languages (German, English, French, Arabic, Farsi, Pashtu, Urdu, Kurmanji)</li> </ul>
Brochure	<ul style="list-style-type: none"> <li>• In 8 languages</li> </ul>
Declaration on data protection	<ul style="list-style-type: none"> <li>• In 8 languages</li> </ul>
Consent to record linkage form <sup>1</sup>	<ul style="list-style-type: none"> <li>• In 8 languages</li> </ul>
Address form and household grid <sup>2</sup>	<ul style="list-style-type: none"> <li>• Electronic form (Mein Kantar)</li> </ul>
Project instruction book	<ul style="list-style-type: none"> <li>• Version for M6</li> </ul>
Other interviewer material	<ul style="list-style-type: none"> <li>• Project description</li> <li>• Contact card (in 8 languages)</li> <li>• HH information card</li> </ul>
Additional interviewer material for M6 to use in the households	All in 8 languages: <ul style="list-style-type: none"> <li>• How is a survey carried out?</li> <li>• FAQ</li> <li>• Aid to fill out the household grid</li> <li>• Postcard for movers</li> <li>• Hygiene flyer</li> </ul>
Additional interviewer material for M6 to use when dealing with shared accommodations	<ul style="list-style-type: none"> <li>• Letter from the Federal Office for Migration and Refugees (BAMF)</li> <li>• Leaflet describing the survey</li> </ul>

<sup>1</sup> Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB Nuremberg): Link to employment history data

<sup>2</sup> Including the so-called "B3 form" used to process address changes.

### Address form and household grid

The address form provided an overview of the household composition as it was last known to Kantar, in the case of sample M6 the form contained the anchor respondent only, due to the household composition being unknown. The interviewers had to document when and with whom the survey was conducted or why a sample member did not participate in the current year. They were also asked to note every single contact attempt made. For sample M6, interviewers did this electronically in the "Mein Kantar" software. Moreover, the interviewers were asked to carefully document any moves of households.

### Project instruction book

In addition to a shorter project description, interviewers in samples M6 also received a much more detailed instruction manual that is about 60 pages long. This manual contained information on special features of the current wave, specific processing instructions, and questionnaires as well as background information on the project.

### Further interviewer materials

Moreover, the interviewers received contact cards for households that could not be reached at home and household information cards with information on anchor-respondents, their name, year of birth, type of questionnaire, and notes.

**Additional interviewer material to use in M6 households**

Because the language barriers in households of refugee samples were often higher than in other samples, interviewers received a few additional laminated sheets in all eight languages to help explain how a survey is carried out to address frequently asked questions and to help fill out the household grid. For movers, interviewers also leaved a postcard with the households that asks them to send their new address to Kantar.

**Additional interviewer material to use when dealing with shared accommodations**

In sample M6, interviewers might have to deal with employees and security personnel at shared refugee accommodations participating households inhabit. For this purpose, the Federal Office for Migration and Refugees (Bundesamt für Migration und Flüchtlinge, BAMF) sent the interviewers a letter and a leaflet describing the survey to be handed over to shared accommodations.

## 4.5 Conducting the survey

### 4.5.1 Survey mode

All questionnaires used in sample M6 were solely available in CAPI mode. However, due to the challenges the Coronavirus pandemic posed to face-to-face fieldwork, a small percentage of interviews (2.7 percent on the household level) were conducted via telephone after an initial personal contact.

### 4.5.2 Fieldwork timings

Face-to-face interviewing for M6 started in mid-August 2020 and ended in February 2021. The fieldwork progress for each month is depicted in **Table 4.6**. Fieldwork was most productive in September and January. In these two months 50.4 percent of the net sample were interviewed.

Table 4.6: **Monthly fieldwork progress M6**

Household Level	Gross Sample		Net Sample	
	Abs.	In %	Abs.	In %
August <sup>1</sup>	101	3.4	55	4.8
September	666	22.2	347	30.4
October	454	15.1	216	18.9
November	263	8.8	100	8.8
December	637	21.2	190	16.7
January	837	27.9	228	20.0
February	42	1.4	5	.4

<sup>1</sup> Including households who refused to take part in the survey prior to start of fieldwork.

## 4.5.3 Translations

Language problems during the interviewing process form a potential major challenge for surveys with populations that recently entered the country as refugees. Although some of the interviewers conducting in the interviews in M3-5 and M6 speak Arabic, Farsi, or Pashtu, it is generally not feasible to match interviewers with special language skills with respondents in such a large, nationwide survey. As implemented successfully in samples M3-5, a bilingual CAPI program was used for boost sample M6. Questions in German and a second language were shown on the screen side by side as shown in **Figure 4.1**. The language was selected at the beginning of the interview.

Figure 4.1: Screenshot bilingual CAPI program – language selection



Table 4.7: Available language versions

	German / English	German / French	German / Arabic	German / Farsi	German / Pashto	German / Urdu	German / Kurmanji
Household questionnaire	✓	✓	✓	✓	✓	✓	✓
Individual + life history questionnaire	✓	✓	✓	✓	✓	✓	✓

**Table 4.8** presents the utilization of each language version for the individual questionnaires. A translated individual questionnaire was used for 2,252 interviews. With 66.8 percent, the German / Arabic version was used most frequently, followed by the German / Farsi version with 17.4 percent and the German / English version at 10.1. The questionnaires in Pashtu, Urdu, and Kurmanji were selected far less frequently.

Table 4.8: Utilization of a certain language versions – individual questionnaire

Total		
	Abs.	In %
German / English	148	12.6
German / French	39	3.3
German / Arabic	838	71.6
German / Farsi	83	7.1
German / Pashto	2	.2
German / Urdu	7	.6
German / Kurmanji	2	.2
Without any translation	51	4.4
<b>Total</b>	<b>1,170</b>	<b>100.0</b>

In addition to the questionnaires and the fieldwork materials that were available in different languages, the interviewers could call a so-called “interpreter hotline” during the process of contacting the households. Then other interviewers that are fluent in either Arabic or Farsi helped the interviewer, e.g., to explain the study background, answer questions and set up an appointment for the interview. But this service could only be used to convince respondents to participate in the study. The interview itself needed to be conducted with the bilingual questionnaires.

## 4.5.4 Panel Maintenance and Incentives

In the first waves of samples M3-5, households did not receive any cash incentives or vouchers. This was due to an assumption that many households still lived in shared accommodations and might experience problems when presented with cash by the interviewers. At the beginning of fieldwork in 2018, interviewers were given a choice between handing over a notebook with a pen as a small gift or bringing a small gift of their choice of less than 5 euros in value (e.g. sweets, small toys for children). After a couple of weeks of fieldwork, some interviewers reported that participants increasingly asked for cash incentives because some had heard that other surveys provide them. Consequently, from 2019 onward interviewers were given the option to incentivize households that answer one household and one individual questionnaire with 15 euros either in cash or in the form of a voucher or gifts. The interviewer was free to choose the form of the incentive. The distribution of these choices across samples M6 is detailed in **Table 4.9**. In 94.7 percent of cases the cash incentive was paid, only 0.4 percent of households received vouchers and or was given a gift as an incentive. Another 0.6 percent of households did not receive any incentive.

Table 4.9: Incentives samples M6

Household level	Abs.	In %
Cash-incentive	1,080	94.7
Voucher	2	0.2
Gift	2	0.2
No Incentive	57	0.6

## 4.5.5 Movers and Tracing

Of the 3,000 households in the administered gross sample of sample M6, 24.5 percent of households were identified as movers and therefore their new address needed to be traced (see **Table 4.10**). What is surprising is that the share of households that needed tracing was much lower than in sample M5 in 2017 (31.4 percent). The tracing was also more successful than it had been in M5 2017: in 49.0 percent of all cases the households could be traced in M6 whereas the success-rate of tracing was a lot lower in sister-sample M5 at 33.1 percent. Most new addresses were gained by Kantar’s supplier for address research, the “Post Address Research”. The interviewers themselves were also very successful, tracing 10.1 percent of households. The inquiry at the local registration offices is not shown to have been of any help tracing households. One needs to keep in mind that “Post Address Research” services include these inquiries and this category is thus missing.

Table 4.10: Movers and sources of new addresses of administered sample of M6

Household Level	Total	
	Abs.	In %
Gross Sample	3,000	100.0
Movers	735	24.5
Success Tracing		
Tracing successful	360	49.0
Tracing not successful	375	51.0
Source		
Interviewer	34	10.1
Postal Service	15	4.4
Participant	0	0.0
Postal Service Address Research	289	85.5



## 4.5.6 Interviewer Characteristics, Training & Monitoring

Because fieldwork in all the refugee samples is carried out by a special group of interviewers, of whom many can speak Arabic, the staff is not only notably different in age compared to the interviewers in samples A-Q. They are also much less experienced as interviewers, as shown in **Table 4.11**. 98.1 percent have less than 5 years of experience working for Kantar compared to only 22.4 percent in samples A-Q.

Table 4.11: Interviewer characteristics

Interviewer Level	Total	
	Abs.	In %
<b>Number of Interviewers</b>	<b>53</b>	<b>100.0</b>
Gender		
Male	47	88.7
Female	6	11.3
Age		
21-39	34	64.2
40-59	17	32.1
60-79	2	3.8
<b>Experience with Kantar</b>		
0-4 years	50	94.3
5-9 years	1	1.9
10-19 years	1	1.9
20-29 years	1	1.9
<b>Experience with SOEP</b>		
0-4 years	52	98.1
5-9 years	1	1.9
Number of Households		
Fewer than 5	5	9.4
5 – 19	9	17.0
20 – 99	28	52.8
More than 100	11	20.8

Kantar places high priority on interviewer monitoring and has put an ISO-certified process in place that is audited regularly. Kantar adheres to the German Business Association of Market and Social Research Institutes (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V., ADM) standards for internal regulation and monitoring of all systems and procedures. This means that a minimum of 10 percent of Kantar’s annual interviews are checked, and every interviewer is monitored at least once a year.

In 2018, we expanded these existing quality control measures for interviewer monitoring in the SOEP projects. The basic interviewer-monitoring concept was expanded to a project-based control system for all SOEP samples: A number of participating households is contacted shortly after the interview by letter or phone asking them to confirm their participation in a regularly conducted interview. In case of inconsistencies and/or irregularities, we attempted to gain clarification through direct contact with respondents, primarily by telephone.

**Table 4.12** shows details for the re-contacting process in sample M6. Of 1,141 households in the net sample, 73.0 percent were re-contacted after the interviews had been conducted in order to assure data quality and identify non-standard behavior by interviewers. In terms of survey mode, phone interviews were the standard approach to ensure a relatively high response rate. Only households for which no telephone number was available received a short paper questionnaire.

Overall, we received feedback from 585 households. This results in a response rate of 70.3 percent overall. Using the results from the re-contacting process as well as analyses of paradata and interview data, we identified one interviewer who had not adhered to our standards in conducting interviews in the 2020 fieldwork period of M6.

Table 4.12: **Interviewer monitoring**

	Total		
	Abs.	In %	In %
By households			
Households in net sample	1,141	100.0	
Re-contacted households	832	73.0	100.0
Households with feedback	585	51.3	70.3
By interviewers			
Interviewers in net sample	50	100.0	
Interviewers with re-contacted households	47	94.0	
Interviewers with non-standard behavior	0	0.0	

## 4.6 Fieldwork results

### 4.6.1 Participation on household level

**Table 4.13** provides a detailed overview on participation in sample M6. In total, 1,141 of 3,000 households in the administered gross sample were interviewed.

Table 4.13: **Participation on household level**

	Total	
	Abs.	In %
Gross Sample	3,000	100.0
<b>Interview (1.0)</b>	<b>1,141</b>	<b>38.0</b>
Complete (1.1)	608	20.3
Partial (1.2)	533	17.8
<b>Non-Interview (2.0)</b>	<b>1,859</b>	<b>62.0</b>
Permanent and soft refusal (2.111, 2.351)	654	21.8
Non-contact (2.20)	692	23.1
Deceased (2.31)	4	0.1
Temporarily physically or mentally unable (2.321)	7	0.2
Permanently physically or mentally unable (2.322)	14	0.5
Language problem (2.331)	55	1.8
Temporarily untraceable (2.4)	10	0.3
Not attempted or worked (3.11)	34	1.1
Moved abroad (4.2)	18	0.6
Household untraceable (4.4)	371	12.4

62.0 percent of the gross sample dropped out of the survey in total. Regardless of the reason, for the first wave of a boost sample all households that drop out will not be eligible for fieldwork in the second wave. Households refusing participation outright and households being untraceable are the most frequent reasons.

**Table 4.14** presents the overall adjusted response rate for sample M6. The overall response rate of all households in samples M6 amounts to 44.3 percent. Compared to the first wave response rates of the most recent refugee boost sample M5 in 2017 (68.8 percent) this is a remarkably low response rate. However, one must keep in mind that SARS-Cov-2 posed a significant challenge for fieldwork in 2020. Households may have been very reluctant to give a face-to-face interview under these circumstances and while in samples A-Q or M1/2 households were given the option to answer PAPI-questionnaires or conduct the interview via telephone, these options were only very restrictively made available for the boost samples in 2020, due to the usual CAPI-only approach being even more important for data quality in boost and refreshment samples.

Table 4.14: **Response Rate at household level**

	Total
Response Rate <sup>1</sup>	<b>44.3</b>

<sup>1</sup> RR= percentage of all households with at least one hh and individual interview and households in the gross sample (gross sample adjusted for households where the anchor respondent is deceased, moved abroad, is permanently untraceable, not attempted or worked).

## 4.6.2 Participation on individual level

**Table 4.15** presents participation on an individual level. Of 1,921 individuals in the gross sample, 63.3 percent were successfully interviewed, while 36.7 percent either dropped out temporarily (8.8 percent) or permanently (27.9 percent). These results must be viewed against the backdrop of an anchor-respondent concept that, which tends to lead to rather low response rates on the individual level, on the one hand and the still ongoing Coronavirus pandemic.

Table 4.15: **Participation on individual level**

	Total	
	Abs.	In %
Gross Sample <sup>1</sup>	1,921	100.0
<b>Interview (1.0)</b>	1,216	63.3
<b>Non-Interview (2.0)</b>	705	36.7
<b>Temporary drop out</b>	170	8.8
Non-contact (2.21)	8	0.4
Temporarily physically or mentally unable/incompetent (2.321)	20	1.0
Soft refusal (2.351)	108	5.6
Others (2.52)	34	1.8
<b>Final Drop out</b>	535	27.9
Permanent Refusal (2.111)	522	27.2
Permanently physically or mentally unable/incompetent (2.322)	9	0.5
Language problem (2.331)	4	0.2

<sup>1</sup> All household members intended to participate in the current wave, restricted to members of participating households.

In the first wave, only household members of participating households are in the gross sample. The response rate on the individual level is provided in **Table 4.16**.

Table 4.16: **Response rate**

	Total
Response Rate	63.3

<sup>1</sup> RR= percentage of all participants in the gross sample of individuals born before 2003.

One major concern for all SOEP samples are the growing partial unit non-response (PUNR) rates. PUNR in sample M6 was very high at 91.1 percent in this year's first wave (**Table 4.17**). PUNR rates in refugee samples tend to be generally higher than in general population samples (first wave M3: 53.7%; first wave M4: 57.3%; first wave M5: 47.7%). This is mainly due to characteristics of the target group. Especially in first waves interviewers report, that it is very hard to convince the refugee households that every member should participate. On the other hand, it is comparatively easy to explain, that a single person was selected from an official register and is asked to participate, which results in the high first wave response rates on household level observed in all previous SOEP refugee samples. Therefore, in the first wave of refugee samples there are exceptionally high PUNR rates. Usually, these rates consolidate somewhat in the coming years. Regarding this year's first wave of M6 the general problem of high PUNR was enormously increased by the Corona pandemic, that occurred during fieldwork and lead to the minimisation of contacts.

Table 4.17: **Partial unit non response**

	Total
PUNR <sup>1</sup>	91.1

<sup>1</sup> Share of households (number of household members > 1) with at least one missing individual questionnaire.

## 4.6.3 Participation by types of questionnaires

**Table 4.18** presents the number of interviews and response rates for each of the two fielded questionnaires. Corresponding with the number of households in the net sample, 1,141 household questionnaires were produced, resulting in an unadjusted response rate of 38 percent. The coverage rate for the individual and life history questionnaire for refugees and Germans/migrants combined is 63.3 percent.

Table 4.18: **Number of interviews and response rate by questionnaire**

	Gross sample <sup>1</sup>	Number of interviews	Coverage rate
Household questionnaire	3,000	1,141	38.0
Individual questionnaire + life history questionnaire module I + II	1,921	1,216	63.3

<sup>1</sup> Gross sample= target population in participating households.

## 4.6.4 Interview length per questionnaire

Median interview lengths for each questionnaire are listed in **Table 4.19**. Completing the individual questionnaire and life-history questionnaire the median was 60 minutes. If the household questionnaire was completed as well, the interviewing time adds up to a total of 77 minutes (median). The interview length might be another contributing factor to high PUNR in this sample specifically but more broadly speaking for the SOEP in general.

Table 4.19: **Median interview length (minutes)**

	Median	Percentiles (5%, 95%)
Household questionnaire	17	(9, 41)
Individual questionnaire + life history questionnaire module for refugees	60	(39, 102)
Individual questionnaire + life history questionnaire module for migrants and Germans	65	(29, 144)

## 4.6.5 Consent to Record Linkage

In all SOEP migration and refugee samples it is an aim to link respondents' survey data with data from the Integrated Employment Biographies Sample (Stichprobe der Integrierten Erwerbsbiografien, IEBS) for as many participants as possible. Based on this, all refugee respondents in sample M6 were asked to give their written consent to this record linkage. As presented in **Table 4.20**, 990 of 1,216 individuals (81.4 percent) gave their consent.

Table 4.20: **Consent to record linkage IEBS**

	Total
Gross Sample <sup>1</sup>	1,216
Consent	990
Consent Rate (in %)	81.4

<sup>1</sup> The gross sample for record linkage in M6 2020 consisted of first-time refugee respondents

In addition to the linkage to the IEBS data, respondents who had told us in this wave or in any previous wave that they had participated in an integration course of the German Federal Office for Migration and Refugees were asked their permission to link their survey data to registration data available in the "Integration Business File" (Integrationsgeschäftsdatei, or InGe) of the BAMF. 77.9 percent of respondents in the gross sample agreed to the linkage, as outlined in **Table 4.21**.

Table 4.21: **Consent to record linkage InGe**

	Total
Gross sample <sup>1</sup>	823
Consent	641
Consent rate (in %)	77.9

## 4.7 Data preparation

Data preparation processes in sample M6 are in line with the processes that are described for samples A-Q in **Section 1.7** of this report.



## 4.8 Delivered data

### Gross Data

Codebook Panel data  
Gross data Household  
Gross data Individuals  
Interviewer data

Codebuch 2020\_M6.pdf  
Hbru\_M6\_2020.dta  
Pbru\_M6\_2020.dta  
Intband\_M6\_2020.sav

### Net Data

Net data Household checked  
Net data Household unchecked  
Net data Individuals + Life history Refugees checked  
Net data Individuals + Life history Refugees unchecked  
Net data Individuals Non-refugees checked  
Net data Individuals Non-refugees unchecked  
Net data Life history Non-refugees checked  
Net data Life history Non-refugees unchecked  
Aggregated net data Refugee Route Module  
Consent to record linkage

H20\_M6.dta  
H20\_M6\_u.dta  
PB20\_M6.dta  
PB20\_M6\_u.dta  
PDM20\_M6.dta  
PDM20\_M6\_u.dta  
LDM20\_M6.dta  
LDM20\_M6\_u.dta  
PB20\_Fluchtrouten\_M6.dta  
EV20\_DIW\_M6.sav

### Other Data

University coding  
Professions, sectors, final coding  
Additional Codes for Individual + Life history Refugee

Hochschul-Abschl\_2020\_M6.dta  
BerufeBranchenAusbildung\_2020\_M6.dta  
Nationen.xlsx  
Zusatzcodes\_2020.xlsx

## 5 Sample M7 / M8a

### 5.1 Introduction

**Table 5.1** gives an overview of the main characteristics of the first waves of the migration boost samples M7 and M8a of 2020. The gross sample for fieldwork consisted of 32,743 households (M7: 19,751; M8a: 12,992). Altogether 1,879 respondents (M7: 783; M8a: 1,096) participated in the survey between July 2020 and February 2021. This results in an adjusted response rate of 11.1 percent for M7 and 15.5 percent for M8a. As is usually the case for boost and refreshment samples in the SOEP, interviews were only administered in CAPI mode. No other modes were used in M7 and M8a, although a small percentage of interviews was conducted via telephone in the CAPI environment. Partial unit non-response (PUNR) was at 84.9 percent for M7 and 81.1 percent for M8a.

Table 5.1: **Summary fieldwork**

	M7	M8a
Fieldwork period	July - February	
Mode (main questionnaires)	CAPI	
Gross sample (hh)	19,751	12,992
Net sample (hh)	783	1,096
Response rate (adjusted; hh) <sup>1</sup>	11.1	15.5
Number of questionnaires	Adults: 3 Youths: 0 Children: 0	
Net sample (individuals)	Adults: 895 Youths: 0 Children: 0	Adults: 1,196 Youths: 0 Children: 0
Questionnaire length (median, in minutes)	Household: 15 Adult: 50	
Partial unit non-response (PUNR)	84.9	81.1

<sup>1</sup> RR = percentage of all households with at least one hh and individual interview in the gross sample (gross sample adjusted for households where the anchor respondent is deceased, moved abroad, is permanently untraceable, not attempted or worked or quality neutral sampling losses).

## 5.2 Background Sample M7 / M8a

In 2020, two boost samples, samples M7 and M8a, were added to the SOEP migration sample system. Like the older migration samples M1 and M2, the Integrated Employment Biographies Sample (IEBS) of the Federal Employment Agency (BA) served as the sampling frame for both boost samples. Boost sample M7's goal was to capture migration dynamics and processes from 2016 to 2018 with a focus on EU migration. To ensure that statistically significant group comparisons can be made, sampling was restricted to the three most significant countries of origin in that time period: Romania, Bulgaria, and Poland. M8a, on the other hand, was designed to help evaluate the skilled worker immigration law (Fachkräfteeinwanderungsgesetz), which came into effect March 1, 2020, and targeted migrants from third countries that came to Germany between 2017 and 2018, sampling them as a control group for a treatment group that will be sampled at a later date.

A total of 43,572 addresses in 125 PSUs were provided from the Integrated Employment Biographies Sample. 22,020 of those belonged to sample M7 and 21,552 belonged to sample M8a. Initially Kantar sampled 14,292 addresses for fieldwork (M7: 7,274; M8a: 7,018). The sampling was stratified proportionally regarding nationality and subsample. As an enormous number of addresses turned out to be not processable (see table 5.13 in chapter 5.6.1), especially in M7, and the response rate among the remaining cases was very low, a second tranche of 18,451 addresses was fielded. In order to compensate the inferior address quality of M7 the second sample consisted of 12,477 M7-addresses and 5,974 M8a-addresses. Therefore, the total gross sample for fieldwork consisted of 32,743 addresses, of which 19,751 were M7-addresses and 12,992 were M8a-addresses.

## 5.3 Structure of the Gross Sample

The expected poor address-quality of addresses sampled from the IEB and the high mobility of the target population resulted in a total sampling of 43,572 addresses. They were distributed over 125 PSUs by the IAB and SOEP-group at the DIW and provided to Kantar. Out of this address-pool, Kantar sampled 32,743 addresses for fieldwork (19,751 in M7 and 12,992 in M8a) (see chapter 5.2).

As presented in **Table 5.2**, most of these households of M7 and M8a are located in the states of North Rhine-Westphalia (19.8 percent), Baden-Wuerttemberg (19.2 percent), and Saarland (15.4 percent). Generally, there were few households in the gross sample located in the eastern states of Germany. More than half of the households (57.4 percent) were positioned in the center of large cities (**Table 5.3**) with either more than 500,000 inhabitants or with 100,000 to 499,999 inhabitants.

Table 5.2: Household characteristics I

Household level	Total		M7		M8a	
	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample</b>	32,743	100.0	19,751	100.0	12,992	100.0
<b>State</b>						
Schleswig-Holstein	737	2.3	499	2.5	238	1.8
Hamburg	1,002	3.1	574	2.9	428	3.3
Lower Saxony	2,694	8.2	1,955	9.9	739	5.7
Bremen	573	1.7	309	1.6	264	2.0
North Rhine-Westphalia	6,478	19.8	4,104	20.8	2,374	18.3
Hesse	3,372	10.3	1,931	9.8	1,441	11.1
Rhineland Palatinate	1,856	5.7	1,348	6.8	508	3.9
Saarland	5,036	15.4	2,867	14.5	2,169	16.7
Baden-Wuerttemberg	6,286	19.2	3,574	18.1	2,712	20.9
Bavaria	267	0.8	177	0.9	90	0.7
Berlin	2,357	7.2	1,069	5.4	1,288	9.9
Brandenburg	559	1.7	329	1.7	230	1.8
Mecklenburg Western Pomerania	285	0.9	158	0.8	127	1.0
Saxony	441	1.3	265	1.3	176	1.4
Saxony-Anhalt	338	1.0	298	1.5	40	0.3
Thuringia	462	1.4	294	1.5	168	1.3

Table 5.3: Household characteristics II

Household level	Total		M7		M8a	
	Abs.	In %	Abs.	In %	Abs.	In %
<b>Gross sample<sup>1</sup></b>	32,740	100.0	19,748	100.0	12,992	100.0
<b>BIK type<sup>2</sup></b>						
0	12,599	38.5	6,625	33.5	5,974	46.0
1	2,588	7.9	1,679	8.5	909	7.0
2	6,187	18.9	3,522	17.8	2,665	20.5
3	3,299	10.1	2,252	11.4	1,047	8.1
4	649	2.0	475	2.4	174	1.3
5	2,229	6.8	1,538	7.8	691	5.3
6	2,524	7.7	1,690	8.6	834	6.4
7	1,767	5.4	1,257	6.4	510	3.9
8	507	1.5	408	2.1	99	0.8
9	391	1.2	302	1.5	89	0.7
<b>Community size<sup>3</sup></b>						
1	1,023	3.1	834	4.2	189	1.5
2	1,672	5.1	1,289	6.5	383	2.9
3	6,106	18.6	3,854	19.5	2,252	17.3
4	5,913	18.1	3,883	19.7	2,030	15.6
5	2,608	8.0	1,548	7.8	1,060	8.2
6	7,453	22.8	4,506	22.8	2,947	22.7
7	7,965	24.3	3,834	19.4	4,131	31.8

<sup>1</sup> For three households this information is not available.

<sup>2</sup> BIK type: 0 (more than 500,000 inhabitants/center), 1 (more than 500,000 inh./periphery), 2 (100,000 to 499,999 inh./center), 3 (100,000 to 499,999 inh./periphery), 4 (50,000 to 99,999 inh./center), 5 (50,000 to 99,999 inh./periphery), 6 (20,000 to 49,999 inh.), 7 (5,000 to 19,999 inh.), 8 (2,000 to 4,999 inh.), 9 (fewer than 2,000 inh.)

<sup>3</sup> Community size: 1 (fewer than 2000 inhabitants), 2 (2,000 to 5,000 inh.), 3 (5,000 to 20,000 inh.), 4 (20,000 to 50,000 inh.), 5 (50,000 to 100,000 inh.), 6 (100,000 to 500,000 inh.), 7 (more than 500,000 inh.).

## 5.4 Questionnaires and fieldwork material

### 5.4.1 Questionnaires

In the first waves of M7 and M8 three questionnaires were fielded: the individual questionnaire for first time respondents (including additional biographical questions) for all adult household members, which had the life-history module integrated in the CAPI-instrument and the household questionnaire for the anchor respondent. In addition to these instruments, anchor-respondents had to answer a short screening questionnaire in order to clarify their membership in the target populations of M7 and M8a respectively. Respondents had to have been born outside of Germany, their stay should not be temporary, and they were to have moved to Germany no earlier than 2016 (M7) or 2017(M8a) respectively. All questionnaires were solely available in CAPI mode. Translation aides were provided only in paper form in four additional languages (see **Section 5.4.3**). With regards to questionnaire content, the household and individual questionnaires were almost identical to the ones used in samples M1/2.

Table 5.4: **Questionnaires and modes sample M7 and M8a**

	CAPI
Screening questionnaire	✓
Household questionnaire	✓
Individual questionnaire + life history questionnaire module	✓

In addition to the questionnaires for respondents, interviewers were asked to complete a short questionnaire about the area the household is located in, the so-called “residential environment questionnaire”.

## 5.4.2 Fieldwork material

In addition to the questionnaires, a whole range of fieldwork materials such as letters, leaflets, and documents for the interviewers were designed, printed, and sent to households and interviewers. **Table 5.5** provides an overview of the different material types that were prepared in sample M7 and M8a. Because the mode was restricted to CAPI, the number of different versions of materials was notably smaller than in samples A-Q. However, many materials were provided in five different languages (German, English, Polish, Romanian, and Bulgarian).

### **Advance letter**

About two weeks before the start of the fieldwork period, households received an advance letter in which the interviewer's visit was announced. It was always sent in German language as well as in one of the four other available languages based on the likelihood of the individual speaking a certain language. The letter included a link to the SOEP website that provided additional information in the different languages.

### **Brochure**

Every household received an eight-page brochure with general information on the survey. The brochures in German and the second language were sent with the advance letter.

### **Declaration on data protection**

Every household got access to a two-page declaration on data protection detailing the organizations that are responsible for processing all respondent data along with a description of data handling and data recipients. The declaration on data protection in German and the second language was sent with the advance letter.

### **Consent to record linkage form**

To all respondents born in or before 2002, a consent to record linkage form was presented that allows an individual linkage between the respondent's data and employment history data available at the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB Nuremberg).

### **Hygiene flyer**

In the face of challenges to face-to-face-interviewing brought on by the Coronavirus pandemic, Kantar implemented strict guidelines for interviewers and households alike to ensure the safety of everyone involved. These guidelines were contained in a flyer that the interviewers brought into the household with them.

Table 5.5: Fieldwork material sample M7/8a

CAPI	
Advance letter	<ul style="list-style-type: none"> <li>• In 5 languages (German, English, Bulgarian, Polish, Romanian)</li> </ul>
Brochure	<ul style="list-style-type: none"> <li>• In 5 languages</li> </ul>
Declaration on data protection	<ul style="list-style-type: none"> <li>• In 5 languages</li> </ul>
Consent to record linkage form <sup>1</sup>	<ul style="list-style-type: none"> <li>• IAB<sup>1</sup></li> </ul>
Address form and household grid <sup>2</sup>	<ul style="list-style-type: none"> <li>• Electronic form (Mein Kantar)</li> </ul>
Project instruction book	<ul style="list-style-type: none"> <li>• Version for M7/M8a</li> </ul>
Other interviewer material	<ul style="list-style-type: none"> <li>• Project description</li> <li>• HH information card</li> <li>• Contact card</li> <li>• Translated paper questionnaires</li> </ul>
Additional interviewer material for M7/8a to use in the households	<ul style="list-style-type: none"> <li>• Postcard for movers</li> <li>• Hygiene flyer</li> </ul>

<sup>1</sup> Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB Nuremberg): Link to employment history data

<sup>2</sup> Including the so-called “B3 form” used to process address changes.

### Address form and household grid

The address form provided an overview of the household composition as it was last known to Kantar, in the case of samples M7 and M8a the form contained the anchor respondent only, due to the household composition being unknown. The interviewers had to document when and with whom the survey was conducted or why a sample member did not participate in the current year. They were also asked to note every single contact attempt made. Interviewers did this electronically in the “Mein Kantar” software. Moreover, the interviewers were asked to carefully document any moves of households.

### Further interviewer materials

Moreover, the interviewers received contact cards for households that could not be reached at home and household information cards with information on anchor-respondent, name, year of birth, type of questionnaire, and notes. For movers, interviewers also leaved a postcard with the households that asks them to send their new address to Kantar.



## 5.5 Conducting the survey

### 5.5.1 Survey mode

All questionnaires used in sample M7 and M8a were solely available in CAPI mode. However, due to the challenges the Coronavirus pandemic posed to face-to-face fieldwork, a small percentage of interviews (5.2 percent in sample M7 and 6.9 percent in sample M8a on the household level) were conducted via telephone after an initial personal contact.

### 5.5.2 Fieldwork timings

Face-to-face interviewing for M7 and M8a started simultaneously in July 2020 and ended in February 2021. The fieldwork progress for each month is depicted in **Table 5.6**. Fieldwork was most productive in November and January. In these two months 54.9 percent of the M7 net sample and 40.6 percent of the M8a net sample were interviewed.

Table 5.6: Monthly fieldwork progress M7/8a

Household Level	Gross Sample				Net Sample			
	M7		M8a		M7		M8a	
	Abs.	In %	Abs.	In %	Abs.	In %	Abs.	In %
July <sup>1</sup>	4,426	22.4	2,689	20.7	77	9.8	186	17.0
August	1,055	5.3	1,622	12.5	53	6.8	186	17.0
September	3,581	18.1	1,521	11.7	37	4.7	55	5.0
October	893	4.5	871	6.7	58	7.4	90	8.2
November	4,199	21.3	1,956	15.1	232	29.6	222	20.3
December	1,624	8.2	1,132	8.7	125	16.0	131	12.0
January	2,878	14.6	2,466	19.0	198	25.3	223	20.3
February	1,095	5.5	735	5.7	3	0.4	3	0.3

<sup>1</sup> Including households who refused to take part in the survey prior to start of fieldwork.

## 5.5.3 Translations

Language problems during the interviewing process form a potential major challenge for surveys with populations that recently entered the country. As implemented successfully in samples M1/2, the household questionnaire as well as the individual and life-history questionnaires were available in four alternative languages other than German, namely English, Bulgarian, Romanian and Polish.

Table 5.7: Translated paper questionnaires

	English	Bulgarian	Romanian	Polish
Household questionnaire	✓	✓	✓	✓
Individual questionnaire	✓	✓	✓	✓
Life history questionnaire	✓	✓	✓	✓
Screening questionnaire	✓	✓	✓	✓

The following **Table 5.8** contains information about language problems and the usage of translated questionnaires in samples M7 and M8a on the individual level. Of all conducted interviews, 50.3 percent required assistance with language problems. This is a very high proportion compared to the older migration samples M1 and M2. This is because M7 and M8a specifically targeted persons who came to Germany relatively recently. For those interviews, translated paper questionnaires were used in the majority of cases. Assistance by a German-speaking person in the same household was relatively infrequent.

Table 5.8: Language problems and usage of translated paper questionnaires

	M7		M8a	
	Abs.	In % (net sample)	Abs.	In % (net sample)
Net sample (individual questionnaire)	895	100.0	1.196	100.0
No language problems occurred/no need for assistance with language problems <sup>1</sup>	296	33.1	743	62.1
Assistance with language problems needed <sup>2</sup>	599	66.9	453	37.9
<b>Of that number:</b>				
German-speaking person in the same household	34	3.8	45	3.8
German-speaking person from outside the household	25	2.8	47	3.9
Professional interpreter	1	0.1	1	0.1
Translated paper questionnaire	574	64.1	371	31.0
<b>Of that number:</b>				
English	22	2.5	371	31.0
Romanian	178	19.9	0	-
Polish	167	18.7	0	-
Bulgarian	207	23.1	0	-

<sup>1</sup> In 26 cases of M7 and 52 cases of M8a there were no answer if translation aids were used or not.

<sup>2</sup> In 38 cases of M7 and 11 cases of M8a two translation aids were used, more than two were used in 2 M7 cases.

## 5.5.4 Panel Maintenance and Incentives

The incentives for respondents in samples M7 and M8a were identical to those in the older migration samples M1 and M2. Adult participants received cash incentives from the interviewers after they completed their interview. The individual questionnaire was rewarded with 10 euros, while the shorter household questionnaire was rewarded with 5 euros. (Table 5.9).

In addition to the individual incentives, interviewers brought a small gift to all households which were presented upon arrival. This year's gift was a high-quality branded shopping bag.

Table 5.9: Incentives samples M7/8a

	M7/8a
Incentives for adults	HH: 5 euros Adult: 10 euros

## 5.5.5 Movers and Tracing

Of the 32,743 households in the gross sample of sample M7 and M8a, in sample M7 43.2 percent of households were identified as movers and therefore their new addresses needed to be traced (see Table 5.10). This goes to show that the addresses drawn from the IEBS were either inaccurate or outdated. This is due to the households consisting of EU migrants from Poland, Bulgaria, and Romania, many of whom are in Germany as seasonal workers and thus are very mobile in Germany and — as EU-citizens — in Europe. For sample M8a, this number is much lower at 26.4 percent. Most new addresses were gained by Kantar's interviewers, who accounted for 28.5 percent of traced households overall. Another significant source for new addresses was Kantar's supplier for address research, the "Post Address Research". The inquiry at the local registration offices is not shown to have been of any help tracing households. One needs to keep in mind that "Post Address Research" services include these inquiries, and this category is thus missing.

Table 5.10: Movers and sources of new addresses of administered sample of M7 and M8a

	M7		M8a	
	Abs.	In %	Abs.	In %
Gross Sample	19,751	100.0	12,992	100.0
Movers	8,572	43.4	3,431	26.4
Success Tracing				
Tracing successful	2,470	28.2	1,713	49.9
Tracing not successful	6,102	71.2	1,718	50.1
Source				
Interviewer	2,234	90.4	1,186	69.2
Postal Service	42	1.7	31	1.8
Participant	4	0.2	10	0.6
Postal Service Address Research	190	7.7	486	28.4

## 5.5.6 Interviewer Characteristics, Training & Monitoring

Information on some interviewer characteristics is collected in **Table 5.11**. In total, 109 interviewers were involved in the first waves of samples M7 and M8a. At 76.1 percent, the share of male interviewers is significantly higher than the share of female ones. The majority of interviewers are between 21 and 39 years old and are responsible for more than 99 households.

**Table 5.11: Interviewer characteristics**

Interviewer level	Total		M7		M8a	
	Abs.	In %	Abs.	In %	Abs.	In %
<b>Number of Interviewers</b>	109	100.0	109	100.0	108	100.0
<b>Gender</b>						
Male	83	76.1	83	76.1	82	75.9
Female	26	23.9	26	23.9	26	24.1
<b>Age<sup>1</sup></b>						
21-39 years	40	36.7	40	36.7	40	37.0
40-59	39	35.8	39	35.8	38	35.2
60-79	30	27.5	30	27.5	30	27.8
<b>Experience with Kantar</b>						
0-4 years of experience with Kantar	77	70.6	77	70.6	77	71.3
5-9 years	16	14.7	16	14.7	16	14.8
10-19 years	13	11.9	13	11.9	12	11.1
20-29 years	2	1.8	2	1.8	2	1.9
30-39 years	1	0.9	1	0.9	1	.9
<b>Experience with SOEP</b>						
0-4 years	80	73.4	80	73.4	80	74.1
5-9 years	21	19.3	21	19.3	21	19.4
10-19 years	7	6.4	7	6.4	6	5.6
20-29 years	1	0.9	1	0.9	1	.9
<b>Number of households (gross)</b>						
Fewer than 5	0	0.0	3	2.8	3	2.8
5 – 19	4	3.7	7	6.4	14	13.0
20 – 99	30	27.5	40	36.7	41	38.0
More than 99	75	68.8	59	54.1	50	46.3

Kantar places high priority on interviewer monitoring and has put an ISO-certified process in place that is audited regularly. Kantar adheres to the German Business Association of Market and Social Research Institutes (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V., ADM) standards for internal regulation and monitoring of all systems and procedures. This means that a minimum of 10 percent of Kantar’s annual interviews are checked, and every interviewer is monitored at least once a year.

In 2018, we expanded these existing quality control measures for interviewer monitoring in the SOEP projects. The basic interviewer-monitoring concept was expanded to a project-based control system for all SOEP samples: A number of participating households is contacted shortly after the interview by letter or phone asking them to confirm their participation in a regularly conducted interview. In case of inconsistencies and/or irregularities, we attempted to gain clarification through direct contact with respondents, primarily by telephone.

**Table 5.12** shows details for the re-contacting process in samples M7 and M8a. Of 1,879 households in the net sample, 62.3 percent were re-contacted after the interviews had been conducted in order to assure data quality and identify non-standard behavior by interviewers. In terms of survey mode, phone interviews were the standard approach to ensure a relatively high response rate. Only households for which no telephone number was available received a short paper questionnaire.

Overall, we received feedback from 658 households. This results in a response rate of 56.2 percent overall. Using the results from the re-contacting process as well as analyses of paradata and interview data, we identified no interviewer who had not adhered to our standards in conducting interviews in the 2020 fieldwork period.

Table 5.12: **Interviewer monitoring**

	Total		
	Abs.	In %	In %
By households			
Households in net sample	1,879	100.0	
Re-contacted households	1,170	62.3	100.0
Households with feedback	658	35.0	56.2
By interviewers			
Interviewers in net sample	103	100.0	
Interviewers with re-contacted households	103	100.0	
Interviewers with non-standard behavior	0	0.0	

## 5.6 Fieldwork results

### 5.6.1 Participation on household level

**Table 5.13** provides a detailed overview on participation in samples M7 and M8a. In total, 1,879 of 32,743 households in the administered gross sample were interviewed.

Table 5.13: **Participation on household level**

	Total		M7		M8a	
	Abs.	In %	Abs.	In %	Abs.	In %
Gross sample	32,743	100.0	19,751	100.0	12,992	100.0
<b>Interview (1.0)</b>	<b>1,879</b>	<b>5.7</b>	<b>783</b>	<b>4.0</b>	<b>1,096</b>	<b>8.4</b>
Complete (1.1)	1,001	3.1	301	1.5	700	5.4
Partial (1.2)	878	2.7	482	2.4	396	3.0
<b>Non-interview (2.0)</b>	<b>30,864</b>	<b>94.3</b>	<b>18,968</b>	<b>96.0</b>	<b>11,896</b>	<b>91.6</b>
Permanent and soft refusal (2.111, 2.351)	4,331	13.2	2,244	11.4	2,087	16.1
Non-contact (2.20)	7,672	23.4	3,960	20.0	3,712	28.6
Deceased (2.31)	41	0.1	31	0.2	10	0.1
Temporarily physically or mentally unable (2.321)	3	0.0	1	0.0	2	0.0
Permanently physically or mentally unable (2.322)	1	0.0	1	0.0	0	0.0
Language problem (2.331)	262	0.8	84	0.4	178	1.4
Not attempted or worked (3.11)	1,788	5.5	1,069	5.4	719	5.5
Not Eligible (4.0)	1,795	5.5	1,087	5.5	708	5.4
Moved abroad (4.2)	3,344	10.2	2,194	11.1	1,150	8.9
Household untraceable (4.4)	11,627	35.5	8,297	42.0	3,330	25.6



94.3 percent of the gross sample dropped out of the survey in total. With 35.5 percent, the most common reason for a drop out was that the respective households could not be traced. The difficulties in processing migrant households are known since the first wave of sample M1 in 2013. However, these difficulties appear to have been exacerbated in boost samples M7 and M8a due to poor address quality and the high mobility of households especially in sample M7, where 11.1 percent of households had already moved abroad when processed. Another 5.5 percent of households were ineligible for fieldwork. Also, difficulties arose with businesses not allowing their employees to participate, interviewers found that some anchor-respondents lived in shared accommodations and that foremen would pocket the cash-incentives and demand they be present during the interview. Unfortunately, these remain anecdotal insights and are as of the time of writing no longer quantifiable due to the restricted access Kantar has to the raw data.

**Table 5.14** presents the overall adjusted response rate for sample M7 and M8a. The overall adjusted response rate of all households in samples M7/8a amounts to 13.3 percent with M7 achieving a lower adjusted response rate (11.1 percent) and M8a a slightly higher one (15.5 percent). These are remarkably low response rates. However, in addition to the poor address quality and the other difficulties that arose during fieldwork, one must keep in mind that SARS-Cov-2 posed a significant challenge for fieldwork in 2020. Households may have been very reluctant to give a face-to-face interview under these circumstances and while in samples A-Q or M1/2 households were given the option to answer PAPI-questionnaires or conduct the interview via telephone, these options were only very restrictively made available for the boost samples in 2020, due to the usual CAPI-only approach being even more important for data quality in boost and refreshment samples.

**Table 5.14: Response Rate at household level**

	Total	M7	M8a
Response Rate <sup>1</sup>	13.3	11.1	15.5

<sup>1</sup> RR= percentage of all households with at least one hh and individual interview and households in the gross sample (gross sample adjusted for households where the anchor respondent is deceased, moved abroad, is permanently untraceable, not attempted or worked).

## 5.6.2 Participation on individual level

**Table 5.15** presents participation on an individual level. Of 3,121 individuals in the gross sample, 67.0 percent were successfully interviewed, while 33.0 percent either dropped out temporarily (21.9 percent) or permanently (11.1 percent). These results must be viewed against the backdrops of an anchor-respondent concept, which tends to lead to rather low response rates on the individual level, on the one hand and the still ongoing Coronavirus pandemic on the other hand. The number of temporary drop outs is relatively high. These respondents will be part of the gross sample again in the next wave.

Table 5.15: **Participation on individual level**

	Total		M7		M8a	
	Abs.	In %	Abs.	In %	Abs.	In %
Gross Sample <sup>1</sup>	3,121	100.0	1,487	100.0	1,634	100.0
<b>Interview (1.0)</b>	2,091	67.0	895	60.2	1,196	73.2
<b>Non-Interview (2.0)</b>	1,030	33.0	592	39.8	438	26.8
<b>Temporary drop out</b>	683	21.9	378	25.4	305	18.7
Non-contact (2.21)	56	1.8	34	2.3	22	1.3
Temporarily physically or mentally unable/incompetent (2.321)	7	.2	6	.4	1	.1
Soft refusal (2.351)	520	16.7	292	19.6	228	14.0
Others (2.52)	100	3.2	46	3.1	54	3.3
<b>Final Drop out</b>	347	11.1	214	14.4	133	8.1
Permanent Refusal (2.111)	299	9.6	190	12.8	109	6.7
Permanently physically or mentally unable/incompetent (2.322)	2	.1	2	.1	0	.0
Language problem (2.331)	46	1.5	22	1.5	24	1.5

<sup>1</sup> All household members intended to participate in the current wave, restricted to members of participating households.

In the first wave, only household members of participating households are in the gross sample. The response rate on the individual level is provided in **Table 5.16**.

Table 5.16: **Response rate**

	Total	M7	M8a
Response Rate	<b>67.0</b>	<b>60.2</b>	<b>73.2</b>

<sup>1</sup> RR= percentage of all participants in the gross sample of individuals born before 2003.

One major concern for all SOEP samples are the growing partial unit non-response (PUNR) rates, which are exceptionally high for boost samples M7 and M8a. PUNR in sample M7 was very high at 84.9 percent in this first wave, PUNR-rates for sample M8a are slightly lower at 81.1 percent (**Table 5.17**). This is due to mostly to the anchor-respondent design and PUNR-rates should consolidate somewhat in the coming years.

Table 5.17: **Partial unit non response**

	Total	M7	M8a
PUNR <sup>1</sup>	<b>83.1</b>	<b>84.9</b>	<b>81.1</b>

<sup>1</sup> Share of households (number of household members > 1) with at least one missing individual questionnaire.

## 5.6.3 Participation by types of questionnaires

**Table 5.18** presents the number of interviews and coverage rates for each of the three fielded questionnaires. Corresponding with the number of households in the net sample, 1,879 household questionnaires were produced, resulting in a coverage rate of 5.7 percent. The coverage rate for the individual questionnaires was 67.0 percent. The coverage rate for the life history questionnaire module was 98.5. The gross sample for the life history questionnaire is equivalent to the number of realized individual questionnaires. The Life history questionnaire module was an additional part at the end of the individual questionnaire.

Table 5.18: **Number of interviews and response rate by questionnaire**

	M7			M8a		
	Gross sample	Number of interviews	Coverage rate	Gross sample	Number of interviews	Coverage rate
Household questionnaire	19,751	783	4.0	12,992	1,096	8.4
Individual questionnaire	1,487	895	60.2	1,634	1,196	73.2
Life history questionnaire module	895	884	98.8	1,196	1,175	98.2

## 5.6.4 Interview length per questionnaire

Median interview lengths for each questionnaire are listed in **Table 5.19**. Completing the individual questionnaire and life-history questionnaire the median was 85 minutes for M7 and 83 minutes for M8a. If the household questionnaire was completed as well, the interviewing time adds up to a total of 101 minutes (median) for M7 and 97 minutes (median) for M8a. The interview length might be another contributing factor to high PUNR in this sample specifically but more broadly speaking for the SOEP in general.

Table 5.19: **Median interview length (minutes)**

	M7		M8a	
	Median	Percentiles (5%, 95%)	Median	Percentiles (5%, 95%)
Household questionnaire	16	(8, 35)	14	(7, 28)
Individual questionnaire	50	(23, 93)	50	(26, 93)
Life history questionnaire module	35	(14, 68)	33	(15, 64)

## 5.6.5 Consent to Record Linkage

In all SOEP migration samples it is an aim to link respondents' survey data with data from the Integrated Employment Biographies Sample (Stichprobe der Integrierten Erwerbsbiografien, IEBS) for as many participants as possible. Based on this, all respondents in sample M7 and M8a were asked to give their written consent to this record linkage. As presented in **Table 5.20**, 1,264 of 2,091 individuals (60.4 percent) gave their consent.

Table 5.20: **Consent to record linkage IEBS**

	M7	M8a
Gross Sample <sup>1</sup>	895	1,196
Consent	515	749
Consent Rate (in %)	57.5	62.6

<sup>1</sup> The gross sample for record linkage in M7 and M8a 2020 consisted of all first-time respondents

## 5.7 Data preparation

Data preparation processes in sample M7 and M8a are in line with the processes that are described for samples A-Q in **Section 1.7** of this report.

## 5.8 Delivered data

### Gross Data

Codebook Panel data	Codebuch_M7-8_2020.pdf
Gross data Household	Hbru_M7-8_2020.dta
Gross data Individuals	Pbru_M7-8_2020.dta
Interviewer data	Intband_M7-8_2020.sav

### Net Data

Net data Household checked	H20_M7-8.dta
Net data Household unchecked	H20_M7-8_u.dta
Net data Individuals checked	P20_M7-8.dta
Net data Individuals unchecked	P20_M7-8_u.dta
Net data life history checked	L20_M7-8.dta
Net data life history unchecked	L20_M7-8_u.dta
Net data screening questionnaire checked	SC20_M7-8.dta
Net data screening questionnaire unchecked	SC20_M7-8_u.dta

### Other Data

University coding	Hochschul-Abschl_2020_M7-8.dta
Professions, sectors, final coding	BerufeBranchenAusbildung_2020_M78.dta
Additional Codes	Nationen.xlsx
	Zusatzcodes_2020.xlsx