ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

DIW Berlin / SOEP (Ed.)

Research Report

SOEP-IS 2019 - PPFAD: Person-related meta-dataset

SOEP Survey Papers, No. 1036

Provided in Cooperation with:
German Institute for Economic Research (DIW Berlin)

Suggested Citation: DIW Berlin / SOEP (Ed.) (2021) : SOEP-IS 2019 - PPFAD: Person-related meta-dataset, SOEP Survey Papers, No. 1036, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:
https://hdl.handle.net/10419/235940

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/

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

1036 2021

SOEP Survey Papers
Series D – Variable Descriptions and Coding

SOEP-IS 2019—PPFAD: Person-related Meta-dataset

SOEP-IS Group

DIW SOEP

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

The aim of the SOEP Survey Papers Series is to thoroughly document the survey's data collection and data processing. The SOEP Survey Papers is comprised of the following series:
Series A – Survey Instruments (Erhebungsinstrumente)
Series B – Survey Reports (Methodenberichte)
Series C – Data Documentation (Datendokumentationen)
Series D – Variable Descriptions and Coding
Series E – SOEPmonitors
Series F – SOEP Newsletters
Series G – General Issues and Teaching Materials

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

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:
SOEP-IS Group, 2021. SOEP-IS 2019—PPFAD: Person-related Meta-dataset. SOEP Survey Papers 1036: Series D – Variable Descriptions and Coding. Berlin: DIW Berlin/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

soeppapers@diw.de

The German Socio Economic Panel at DIW Berlin

SOEP-IS 2019–PPFAD: Person-related Meta-dataset

SOEP-IS Group

2021

The file ppfad is part of a collection, which is released with doi:10.5684/soep.is.2019.

Contents

1 General Information 3

2 Variables in File PPFAD 3
cid – Original Household Number . . . 3
pid – Never Changing Person ID . . . 3
psample – Sample Member . . . 3
sex – Sex . . . 4
gebjahr – Year of Birth . . . 4
gebmonat – Month of Birth . . . 5
gebmoval – Month of Birth, Data Source . . . 5
erstbefr – Year First Surveyed . . . 6
letztbef – Year of Last Survey . . . 6
eintritt – Year First Contacted . . . 7
austritt – Year of Last Contact . . . 8
todjahr – Year Died, 4 Digits . . . 8
todinfo – Year Died, Information Source . . . 9
loc1989 – Where did you live in 1989? . . . 9
germborn – Born In Germany . . . 10
immiyear – Year of Immigration to Germany . . . 10
corigin – Country of Origin . . . 11
migback – Migration Background . . . 12
miginfo – Information Source of MIGBACK . . . 13
netto$ – Current Wave Survey Status [in year] [generic] . . . 13
pop$ – Sample Membership [in year] [generic] . . . 14
hid$ – Current Wave Household Number [in year] [generic] . . . 15
sampreg$ – Current Sample Region [in year] [generic] . . . 15
casemat$ – Case Match, Combined Panel Households [in year] [generic] . . . 15

1 General Information

This file is designed to support longitudinal analysis linking personal information from various waves of SOEP-IS. Each person saved in at least one $PBRUTTO file is also in PPFAD. The only sorting key (SORT-ID) is the unchanging person identification number PID. The variable CID contains the unchanging original household number whereas the variables HID$$ contain the current household number of the particular household in which the person lived and was interviewed at the time of the survey, either as an adult household member or child. The variables NETTO$$ give the wave-specific interview status or tell which file contains a person in a particular wave ($P, $KIND, etc.) or whether the person has left SOEP-IS. The variables $HID and $NETTO are coded "–2" for those waves in which a person is not yet or no longer part of the SOEP survey population. In addition, PPFAD contains basic demographic variables, which have been longitudinally tested to some extent:
o SEX (sex)
o GEBJAHR (4-digit birth year)
o TODJAHR (4-digit year of death, if applicable)
o IMMIYEAR (year of first immigration to Germany)
o GERMBORN (born in Germany)
o CORIGIN (country of origin)
o MIGBACK (migration background)
o MIGINFO (information source for MIGBACK)
It is recommended that these tested variables are used for cross-sectional and longitudinal analysis. These variables are adjusted on a wave-by-wave basis in the framework of demographic testing. Therefore, a small number of individuals from different original SOEP-IS households moved into the same household. These individuals share the same current household number for the current wave (HID) but have different original household numbers, CID. With the variable $CASEMAT, these households are assigned the value of the CASE-ID (CID) that also needs to be taken into account. All other households are assigned the value -2.
Further methodological variables (year, four-digit):
o EINTRITT: year a person joined SOEP-IS
o AUSTRITT: year a person ultimately left SOEP-IS
o ERSTBEFR: year of a person's first interview
o LETZTBEF: year of a person's most recent interview

2 Variables in File PPFAD

cid – Original Household Number

pid – Never Changing Person ID

psample – Sample Member

? Please choose a sample! *(from: soep-is/soep-is-2014-f/Q2;sample[2245])*

5	[5] E Refreshment 1998	1342
9	[9] I Incentivation 2009	4040
22	[22] S1 Supplementary 2012	2607
25	[25] S2 Supplementary 2013	2716
26	[26] S3 Supplementary 2014	1989
28	[28] Bonn Intervention Panel 2014	2246

32	[32] S4 Supplementary 2016	2375
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

Unchanging sample information

sex – Sex

? Sex: *(from: soep-is/soep-is-2019/PA1:lsex[28037])*

1	[1] Male	8507
2	[2] Female	8778
-1	[-1] No Answer	30
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

Respondent's sex, longitudinally tested.

gebjahr – Year of Birth

? Your date of birth: // 3 Year *(from: soep-is/soep-is-2019/PA2:lgeb[28038])*

1907		1
1908		2
1909		1
1910		1
1913		1
1914		1
1915		1
1917		4
1918		3
1919		11
1920		6
1921		17
1922		19
1923		22
1924		25
...	(87 rows omitted)	16338
2012		124
2013		110
2014		97
2015		78
2016		86
2017		85

2018		46
2019		47
2020		1
-1	[-1] No Answer	188
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

Respondent's year of birth, longitudinally tested.

gebmonat – Month of Birth

? Your date of birth: // 2 Month *(from: soep-is/soep-is-2019/PA2:lgebmo[28038])*

1	[1] January	1209
2	[2] February	1139
3	[3] March	1310
4	[4] April	1200
5	[5] May	1234
6	[6] June	1154
7	[7] July	1250
8	[8] August	1226
9	[9] September	1188
10	[10] October	1211
11	[11] November	1088
12	[12] December	1107
-1	[-1] No Answer	2999
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

gebmoval – Month of Birth, Data Source

1	[1] Generated from data set pbrutto	12
2	[2] Reported by mother data set H	1
3	[3] Generated from data set P	14
4	[4] Info from data set biobirth/biobrthm	429
5	[5] Info from data set OLDER THAN 48 MONTH	1866
6	[6] Info from data set 24 TO 47 MONTH	222
7	[7] Info from data set 0 TO 23 MONTH	529
8	[8] Info from data set BIOAGE08b	0
9	[9] Info from data set BIOAGE08a	13
10	[10] Info from data set BIOAGE06	27
11	[11] Info from data set BIOAGE03	26
12	[12] Info from data set BIOAGE01	108
13	[13] Info from data set YOUTH	120

14	[14] Info from data set S	775
15	[15] Info from data set LELA	10174
-1	[-1] No Answer	2999
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

GEBMOVAL: Data source on month of birth.

erstbefr – Year First Surveyed

1998		724
1999		38
2000		22
2001		25
2002		27
2003		33
2004		15
2005		25
2006		22
2007		30
2008		15
2009		2451
2010		273
2011		101
2012		1711
2013		2000
2014		2302
2015		255
2016		1730
2017		180
2018		174
2019		118
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	5044
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

letztbef – Year of Last Survey

1998	6
1999	14
2000	17
2001	16
2002	13
2003	15

2004		15
2005		18
2006		29
2007		15
2008		12
2009		668
2010		378
2011		505
2012		568
2013		738
2014		915
2015		582
2016		1050
2017		772
2018		942
2019		4983
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	5044
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

eintritt – Year First Contacted

1998		936
1999		27
2000		16
2001		27
2002		20
2003		30
2004		25
2005		18
2006		28
2007		25
2008		13
2009		3454
2010		296
2011		96
2012		2391
2013		2606
2014		4096
2015		213
2016		2463
2017		232
2018		164
2019		139
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	0

-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

austritt – Year of Last Contact

1999		3
2000		13
2001		20
2002		13
2003		15
2004		20
2005		11
2006		23
2007		31
2008		27
2009		18
2010		709
2011		648
2012		547
2013		743
2014		945
2015		1324
2016		819
2017		1163
2018		1144
2019		9079
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

todjahr – Year Died, 4 Digits

2011		16
2012		34
2013		37
2014		44
2015		59
2016		53
2017		65
2018		56
2019		59
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	16892
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0

-6	[-6] Questionnaire Version With Modified Filtering	0

todinfo – Year Died, Information Source

1	[1] From Annual Survey	423
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	16892
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

For all persons who have been identified as deceased over the course of SOEP, the variable TODINFO gives the source of this information. The variable TODJAHR contains the four-digit year of death or a missing value code:
(-2): persons, for whom it is unknown whether they are deceased (that is, both persons still living up to that wave, and persons whose exact whereabouts is unknown and have dropped out of SOEP-IS)
(-1): persons, for whom the fact of death is known, but the year of death is unknown.
$$$$: factual year of death for persons whose death could be firmly established
The information is taken from the annual survey or/and the question about family events in the personal questionnaire.

loc1989 – Where did you live in 1989?

? Where did you live before German reunification, i.e. before 1989? *(from: soep-is/soep-is-2019/Q210:lwied[28242])*

1	[1] East Germany (GDR) incl. East Berlin	1958
2	[2] West Germany (FRG) incl. West Berlin	6901
3	[3] Abroad (Ausland)	942
4	[4] Born 1989 or later	5390
-1	[-1] No Answer	84
-2	[-2] Does Not Apply	1976
-3	[-3] Answer Improbable	64
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable's information is mainly based upon the corresponding question in the biography questionnaire. Additionally, persons who ansered 'East Germany' or 'West Germany' but stated that they immigrated after 1989 were corrected to '[3] Abroad'. Persons who initially answered 'Abroad' but also stated that they immigrated to Germany before 1989 were corrected to '[-3] Answer Improbable'.
For individuals with the values '[-1] No Answer' or '[-2] Does Not Apply' no valid information could be generated. Note that '[-2] Does Not Apply' is assigned to adults who never answered an individual questionnaire. Persons with NETTO$ codes between 210 and 229 are coded to '[-1] No Answer' if relevant information is missing.

germborn – Born In Germany

? Were you born in Germany? *(from: soep-is/soep-is-2019/Q176:lsta1[28190])*

1	[1] Born in Germany or immigr. <1949	13672
2	[2] Immigr. >1948	1617
-1	[-1] No Answer	112
-2	[-2] Does Not Apply	1914
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable's information is mainly based upon the corresponding question in the biography questionnaire. Additionally, persons who stated that they immigrated to Germany before 1949 were automatically assigned the value '[1] Immigrant Before 49 Or German Born'. For children (NETTO$ codes between 220 and 229) their parents' information about birthplace and year of immigration was used to generate the variable whereby single parentship was taken into consideration as well.
For individuals with the values '[-1] No Answer' or '[-2] Does Not Apply' no valid information could be generated. Note that '[-2] Does Not Apply' is assigned to adults who never answered an individual questionnaire. Persons with NETTO$ codes between 210 and 229 are coded to '[-1] No Answer' if relevant information is missing.

immiyear – Year of Immigration to Germany

? When did you move to Germany? // Year: *(from: soep-is/soep-is-2019/Q178:lzug01[28192])*

1949		3
1950		1
1951		1
1952		2
1953		1
1954		2
1955		2
1956		4
1957		8
1958		4
1959		3
1960		5
1961		7
1962		3
1963		5
...	(47 rows omitted)	1378
2011		26
2012		22
2013		30
2014		19
2015		22
2016		5
2017		2

2018		1
2019		2
-1	[-1] No Answer	1510
-2	[-2] Does Not Apply	14025
-3	[-3] Answer Improbable	222
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable's information is mainly based upon the corresponding question in the biography questionnaire. Additionally, persons who stated that they were born in Germany were assigned the value '[-2] Does Not Apply' as well as those stating that they immigrated to Germany before 1949. For children (NETTO$ codes between 220 and 229) their parents' information about birthplace and year of immigration was used to generate the variable whereby single parentship was taken into consideration as well.
For individuals with the value '[-1] No Answer' no valid information could be generated. Note that '[-1] No Answer' in immiyear includes—other than in germborn—persons with all NETTO$ codes. Due to '[-2] Does Not Apply' meaning that no immigration occured no distinction between those with invalid NETTO$ codes and those with real missing information could be made for immiyear. '[-3] Answer Improbable' reports a conflict between information from different biography variables.

corigin – Country of Origin

? In which country were you born? *(from: soep-is/soep-is-2019/Q177:lsta2[28191])*

1	[1] Germany	13673
2	[2] Turkey	198
3	[3] Ex-Yugoslavia	15
4	[4] Greece	26
5	[5] Italy	59
6	[6] Spain	8
7	[7] Ex-GDR (Country Of Origin Only)	2
10	[10] Austria	39
11	[11] France	21
12	[12] Benelux	0
13	[13] Denmark	3
14	[14] Great Britain	17
15	[15] Sweden	1
16	[16] Norway	0
17	[17] Finland	6
...	(156 rows omitted)	1216
174	[174] Madagascar	1
175	[175] Grenada	0
176	[176] Lesotho	0
177	[177] Bhutan	1
178	[178] Rwanda	0
179	[179] Malawi	0
180	[180] Bessarabia	0
222	[222] Eastern Europe	0
333	[333] unspec. foreign country	0

-1	[-1] No Answer	111
-2	[-2] Does Not Apply	1696
-3	[-3] Answer Improbable	222
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable's information is mainly based upon the corresponding question in the biography questionnaire. Additionally, persons with NETTO$ codes between 210 and 229 and GERMBORN equal to '[1] Immigrant Before 49 Or German Born' are assigned a '[1] Germany' in CORIGIN. If information about the place of birth is missing and GERMBORN is equal to '[2] Immigrant > 1948' valid information about each person's nationality is used to generate CORIGIN.
For individuals with the values '[-1] No Answer' or '[-2] Does Not Apply' no valid information could be generated. Note that '[-2] Does Not Apply' is assigned to adults who never answered an individual questionnaire. Persons with NETTO$ codes between 210 and 229 are coded to '[-1] No Answer' if relevant information is missing. '[-3] Answer Improbable' reports a conflict between information from different biography variables.

migback – Migration Background

? Were you born in Germany? *(from: soep-is/soep-is-2019/Q176:lsta1[28190])*
? Do you have German citizenship? *(from: soep-is/soep-is-2019/Q179:lsta1a[28193])*

1	[1] No Migration Background	10029
2	[2] Direct Migration Background	1617
3	[3] Indirect Migration Background	2054
4	[4] Migration Background, Not Differentiable	149
-1	[-1] No Answer	1761
-2	[-2] Does Not Apply	1705
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

MIGBACK provides information on migration background. It considers the nationality of the individual, his/her migration history as well as that of his/her parents. In order to do so information on country of birth, citizenship and its change over time (in case of naturalization for example), and if available, parental information (direct information or proxy information) are combined. Direct/Indirect migration background indicates whether the person immigrated by him/herself (direct migration background) or if she/he is of migrant origin but born in Germany (indirect migration background). That applies for children with parents who have an indirect migration background. Respondents assigned the code 4 have a migration background. As the information on the region of birth is missing, it is not possible to know whether they migrated by themselves of if they were born in Germany.
Note that code '[1] No Migration Background' also includes persons of German nationality born in Germany whose parental information are not available. Some of them might be migrants' offspring. Thus, the MIGBACK variable slightly underestimates the number of persons belonging to code '[3] Indirect Migration Background'.
Note that in principle, an individual who gains a foreign nationality will be classified as being of migrant origin even if this was not the case before. Thus, the migration background

variable as applied in SOEP-IS takes a retrospective perspective and will be updated every year.
Ethnic German repatriates ("Spät-/Aussiedler") in this classification are coded '[2] Direct Migration Background'. Foreigners who were born in Germany and naturalized are coded '[3] Indirect Migration Background'.
In order to provide the highest possible transparency we include a variable informing about the sources used to create the migration background variable. This information is provided in MIGINFO.
For individuals with the values '[-1] No Answer' or '[-2] Does Not Apply' no valid information could be generated. Note that '[-2] Does Not Apply' is assigned to adults who never answered an individual questionnaire. Persons with NETTO$ codes between 210 and 229 are coded to '[-1] No Answer' if relevant information is missing.

miginfo – Information Source of MIGBACK

? Were you born in Germany? *(from: soep-is/soep-is-2019/Q176:lsta1[28190])*
? Do you have German citizenship? *(from: soep-is/soep-is-2019/Q179:lsta1a[28193])*

1	[1] Direct Personal w/o Parental Info	8567
2	[2] Proxy Personal w/o Parental Info	23
3	[3] Direct Personal With Parental Info	891
4	[4] Proxy Personal With Parental Info	4368
5	[5] Only Parental Info And Place Of Birth	0
-1	[-1] No Answer	1761
-2	[-2] Does Not Apply	1705
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

The MIGINFO variable informs on the information used to generate MIGBACK. There are two distinctions between used information. Firstly, the personal information. Either it stems from the interviewee or it is reported by someone else. Secondly, there is a further differentiation between persons for which we have parental information and those whose parents were not interviewed.
Direct information stem from the respondent (NATION$, GERMBORN or variables about citizenship acquisition in $P). A person's migration background is based on proxy information if at least one used information is either reported by the parents or by the interviewer (PNAT$ out of $PBRUTTO or the respective Infratest information).
For individuals with the values '[-1] No Answer' or '[-2] Does Not Apply' no valid information could be generated. Note that '[-2] Does Not Apply' is assigned to adults who never answered an individual questionnaire. Persons with NETTO$ codes between 210 and 229 are coded to '[-1] No Answer' if relevant information is missing.

netto$ – Current Wave Survey Status [in year] [generic]

210	[210] Interviewee With Succesful Interview (_P)	713
212	[212] Individual Questionnaire And Person Biography	0
213	[213] Individual Questionnaire And Youth Biography	0
214	[214] Individual Questionnaire And Other Questionnaires	0

215	[215] Individual Questionnaire And Experiments, Test	0
216	[216] Individual Questionnaire, First Time Surveyed, 17 Years	11
217	[217] Youth Biography First Time Surveyed, 17 Years	0
218	[218] Individual Questionnaire And Child und 17 Years	0
219	[219] Individual Questionnaire Without Household Interview	0
220	[220] Children in Succesfully Interviewed Households (_Kind)	187
221	[221] Children With Mother-Child Questionnaire_I, 0-1 Year	0
222	[222] Children With Mother-Child Questionnaire_II, 2-3 Year	0
223	[223] Children With Mother-Child Questionnaire_III, 5-6 Year	0
224	[224] Parents with children 7-8 year	0
225	[225] Parents with children 9-10 year	0
...	(13 rows omitted)	25
289	[289] Repatriate - (was drop out [90])	0
290	[290] Individual Dropouts PBR_EXIT	0
291	[291] Moved abroad	0
292	[292] Moved abroad (abroad)	0
293	[293] Moved abroad (exit)	0
294	[294] Person Gap with advices	0
297	[297] advice to dead person (exit)	0
298	[298] advice to dead person (_VP)	0
299	[299] Died	0
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	16379
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

NETTO$ follows the SOEP logic in principle, although some of the provided value labels are not relevant for SOEP-IS at the moment. The main difference is that SOEP-IS uses 3-digit numbers and features an integrated personal interview, capturing the biography information as well as the personal information (code 210). If a person does not fill out the questionnaire, the code is set 230. [This information can be related to a specific variable and is not necessarily generic.]

pop$ – Sample Membership [in year] [generic]

1	[1] Private HH, German HH-Head	911
2	[2] Private HH, Foreign HH-Head	25
3	[3] Institutional. HH, German HH-Head	0
4	[4] Institutional. HH, Foreign HH-Head	0
5	[5] Not Compl. Private HH, German HH-Head	0
6	[6] Not Compl. Private HH, Foreign HH-Head	0
7	[7] Not Compl. Institutional. HH, German HH-Head	0
8	[8] Not Compl. Institutional. HH, Foreign HH-Head	0
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	16379
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

POP$ was derived from $WUM2 (see private household, institutional household in $HBRUTTO) as well as $PNAT and $STELL (nationality and relationship to head of household in $PBRUTTO). Missing values were imputed based on the person's history. Thus, the only admissible missing value is –2, meaning not applicable.
This variable is therefore particularly important, as it enters into the determination of cross-sectional weights. [This information can be related to a specific variable and is not necessarily generic.]

hid$ – Current Wave Household Number [in year] [generic]

Identifier for the household (HHNRAKT) in which a person lived at the time of the interview. [This information can be related to a specific variable and is not necessarily generic.]

sampreg$ – Current Sample Region [in year] [generic]

1	[1] West-Germany	767
2	[2] East-Germany	169
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	16379
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

Households have been identified as living in either in the West or the East sample region. This information is time-dependent; thus for each year, there is the wave-specific information SAMPREG$.
For analyses differentiating between East and West Germany, the current sample region variable SAMPREG$ should be used, and not the time-independent original sample variable PSAMPLE. [This information can be related to a specific variable and is not necessarily generic.]

casemat$ – Case Match, Combined Panel Households [in year] [generic]

0	[0] CASE With HH Details	0
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	17315
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

For Wave 2010 of SOEP-IS, we had two individuals from different original households (CASES) living in one common household. In this case, people with identical values for HID in one wave have different values for CID (CASE-ID).
The CASEMAT$ variable is assigned a value only for those persons living in the households described above, and contains the CID of the other household members.
This information is not relevant when linking person and household data based on the current household number HID in non-hierarchically structured databanks. [This information can be related to a specific variable and is not necessarily generic.]