1034 2021

SOEP Survey Papers
Series D – Variable Descriptions and Coding

SOEP-IS 2019—PGEN: Person-related Status and Generated Variables

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

Please cite this paper as follows:
SOEP-IS Group, 2021. SOEP-IS 2019—PGEN: Person-related Status and Generated Variables. SOEP Survey Papers 1034: Series D – Variable Descriptions and Coding. Berlin: DIW Berlin/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–PGEN: Person-related Status and Generated Variables

SOEP-IS Group

2021

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

Contents

1 Overview 4

2 Variables in file pgen 4

cid – Case-ID, Original Household Number . . . 4
hid – Current Wave HH Number . . . 4
pid – Never Changing Person ID . . . 4
syear – Survey Year . . . 4
pgerwtyp – Type of occupation . . . 4
pgerljob – Working in Occupation Trained for . . . 5
pgbetr – Size of the Company . . . 6
pgoeffd – Civil Service . . . 6
pgausb – Required Training for Job . . . 7
pgpartz – Partner Identicator . . . 7
pgpartnr – Person ID number of partner . . . 8
pgnation – Citizenship - nationality . . . 9
pgsbil – Diplomas/degrees from secondary/tertiary . . . 9
pgbbil01 – Vocational degree attained . . . 10
pgbbil02 – Completed college education . . . 10
pgbbil03 – No vocational degree . . . 11
pgsbila – Secondary school degrees/diplomas abroad . . . 11
pgbbila – Occupational Training in abroad . . . 12
pgsbilo – Secondary school degree/diploma - East Germany . . . 12
pgbbilo – Vocational degree attained - East Germany . . . 13
pgfamstd – Marital status in survey year . . . 13
pgbilzt – Amount of education or training (in years) . . . 14
pgerwzt – Length Of Time With Firm . . . 15
pgtatzt – Actual Weekly Work Time . . . 16
pgvebzt – Agreed Upon Weekly Work Time . . . 17
pguebstd – Overtime per Week . . . 18
pglfs – Labor Force Status . . . 19
pgis88 – Current Occupational Classification (ISCO-88 Com) . . . 21
pgis08 – Current Occupational Classification (ISCO-08) . . . 23
pgisei88 – ISEI-Status following Ganzeboom (based on IS88) . . . 24
pgisei08 – ISEI-Status following Ganzeboom (based on IS08) . . . 25
pgmps92 – Magnitude Prestige Scale (based on KldB92) . . . 25
pgmps08 – Magnitude Prestige Scale (based on IS88 recoded from IS08) . . . 26
pgnace – Two-digit NACE Industry-Sector (NACE Rev. 1.1, Sector) . . . 27
pgnace2 – Two-digit NACE Industry-Sector (NACE Rev. 2, Sector) . . . 28
pgsiops88 – TREIMANS STANDARD INT.OCC.PR.SCORE (based on IS88) . . . 29
pgsiops08 – TREIMANS STANDARD INT.OCC.PR.SCORE (based on IS08) . . . 30
pgegp88 – ERIKSON and GOLDTHORPE Class Category (based on IS88) . . . 31
pgegp08 – ERIKSON and GOLDTHORPE Class Category (based on IS88 recoded from IS08) 32
pgklas92 – Current Occupational Classification (KldB92) . . . 33
pgklas10 – Current Occupational Classification (KldB2010) . . . 35
pgautono – Autonomy in occupational activity . . . 36
pgisced – Highest degree/diploma attained, ISCED-1997 . . . 37
pgcasmin – Highest degree/diploma according to CASMIN . . . 38
pgstib – Occupational Position . . . 39
pgmonth – Month of interview . . . 40

pgmode – Interview method . . . 41
pglabgro – Current gross labor income in euros (generated) . . . 41
pgi1labgro – 1. Imput. Akt. Bruttoerwerbseink.(gen) in Euro [1/5] . . . 42
pgi2labgro – 2. Imput. Akt. Bruttoerwerbseink.(gen) in Euro [2/5] . . . 43
pgi3labgro – 3. Imput. Akt. Bruttoerwerbseink.(gen) in Euro [3/5] . . . 44
pgi4labgro – 4. Imput. Akt. Bruttoerwerbseink.(gen) in Euro [4/5] . . . 45
pgi5labgro – 5. Imput. Akt. Bruttoerwerbseink.(gen) in Euro [5/5] . . . 46
pgimpgro – Imputation flag for LABGROxx . . . 47
pglabnet – Current net labor income (generated) in euros . . . 47
pgi1labnet – 1. Imput. Akt. Nettoerwerbseink.(gen) in Euro [1/5] . . . 48
pgi2labnet – 2. Imput. Akt. Nettoerwerbseink.(gen) in Euro [2/5] . . . 49
pgi3labnet – 3. Imput. Akt. Nettoerwerbseink.(gen) in Euro [3/5] . . . 50
pgi4labnet – 4. Imput. Akt. Nettoerwerbseink.(gen) in Euro [4/5] . . . 50
pgi5labnet – 5. Imput. Akt. Nettoerwerbseink.(gen) in Euro [5/5] . . . 51
pgimpnet – Imputation flag for LABNETxx . . . 52
pgallbet – Core size category of the company . . . 53
pgemplst – Employment status . . . 53
pgexpft – Working experience full-time employment . . . 54
pgexppt – Working experience part-time employment . . . 55
pgexpue – Unemployment experience . . . 56
pgjobch – Occupational Change . . . 57
pgfield – Field of tertiary education . . . 58
pgdegree – Type of tertiary degree . . . 59
pgtraina – Apprenticeship, two-digit occupation KldB92 . . . 61
pgtrainb – Vocational school, two-digit occupation KldB92 . . . 62
pgtrainc – Higher vocational school, two-digit occupation KldB92 . . . 63
pgtraind – Civil servant training, two-digit occupation KldB92 . . . 64
pgfdt_f – Data source FIELD, DEGREE, TRAIN . . . 65
pgbilztch – Change in Education since last survey / last year . . . 65
pgbilztev – Change in Education, total observed period . . . 65
pgsndjob – Current gross secondary income in euros . . . 66
pgimpsnd – Imputation flag for SNDJOB . . . 67

1 Overview

Variables in the file pgen documented here are generated mostly from the answers in the personal questionnaire. There is one row for each wave (syear) a person (pid) participated in the survey.

2 Variables in file pgen

cid – Case-ID, Original Household Number

hid – Current Wave HH Number

pid – Never Changing Person ID

syear – Survey Year

1998	1998	724
1999	1999	750
2000	2000	755
2001	2001	766
2002	2002	780
2003	2003	795
2004	2004	792
2005	2005	799
2006	2006	797
2007	2007	797
2008	2008	794
2009	2009	3226
2010	2010	2745
2011	2011	2506
2012	2012	3696
2013	2013	5141
2014	2014	6638
2015	2015	5897
2016	2016	7097
2017	2017	6179
2018	2018	5633
2019	2019	4983
-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

pgerwtyp – Type of occupation

? Are you currently employed? Which one of the following applies best to your status? *(from: soep-is/soep-is-2019/Q302:perw[28325])*

1	[1] Not Employed, Green	27145
2	[2] Not Employed (First Surveyed) Not Applicable Since 94	0

3	[3] Employed (First Surveyed) Not Applicable Since 94	0
4	[4] Empl. Exc Change	24139
5	[5] Empl. No Info If Change	3816
6	[6] Empl. With Change, Also First Time Employment	6971
7	[7] Empl. With Near-Retirement Part-time	219
-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

This variable is generated from the question on whether a respondent has changed jobs since the beginning of the previous year, which is a central filter variable in the questionnaire.
Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
An alternative variable is PGJOBCH (see below), which is an improved version of PGERWTYP, as it is generated in a longitudinally consistent way and contains an additional category for first-time employed persons.
Respondents from the supplementary samples are not being asked about the information on job change; hence, in the year when these samples enter the SOEP-IS, the majority of the employed persons fall into the category [5] (Employed, no info if change).

pgerljob – Working in Occupation Trained for

1	[1] Yes	4660
2	[2] no	2593
3	[3] In Education	412
4	[4] has No Job Training	397
-1	[-1] No Answer	354
-2	[-2] Does Not Apply	6909
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	46965
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable is designed to offer annual data on all employed persons, indicating whether they are working in the occupation they were trained for.
Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
Because detailed information on working in occupation trained for is not assessed in the Questionnaire of the SOEP Innovation Sample, PGERLJOB is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGERLJOB is coded to "-5" (not contained in questionnaire).

pgbetr – Size of the Company

? How many people does the company employ as a whole? This does not refer to a local unit of the company, but to the entire company. *(from: soep-is/soep-is-2019/Q427:pgesunt[28443])*

1	[1] LT 5	1960
2	[2] GE 5 LT 10	2249
3	[3] GE 11 LT 20	2096
4	[4] Until 90: LT 20	0
5	[5] 91-04: GE 5 LT 20	553
6	[6] GE 20 and LT 100	5470
7	[7] From 100 To Les Than 200	2610
8	[8] Until 98: GE 20 LT 200	229
9	[9] 200 Up To 2000	6250
10	[10] 2000 And More	8789
11	[11] Self-Employed Without Coworkers	0
12	[12] Do not know	0
-1	[-1] No Answer	1308
-2	[-2] Does Not Apply	30770
-3	[-3] Answer Improbable	6
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable is designed to offer annual data on company size for all employed persons. Please pay attention to special codes 4, 5, and 8! These codes were necessary due to the differentiation of items for small and medium-sized companies over the years.
Not all employed persons are asked the question on firm size on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
Please also see PGALLBET for a broader categorization of the firm size, which is appropriate for analyses that include all sample years.
Self-employed are not included in this variable. Detailed information about the company size of self-employed is included in the variable PGSTIB.

pgoeffd – Civil Service

? Does the company in which you are employed belong to the public sector? *(from: soep-is/soep-is-2019/Q314:poed[28334])*

1	[1] Yes	8251
2	[2] no	26317
-1	[-1] No Answer	559
-2	[-2] Does Not Apply	27163
-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 is designed to provide annual data on employment in the civil service for all employed persons.

Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.

pgausb – Required Training for Job

1	[1] Yes	63
2	[2] Intro. To Job	1224
3	[3] On-The-Job Training	482
4	[4] Courses	232
5	[5] Vocational Training	4187
6	[6] Technical School, Engineering (East) 90-96	0
7	[7] Technical College, University until 1998	82
8	[8] Technical College since 1999	767
9	[9] University since 1999	1009
-1	[-1] No Answer	370
-2	[-2] Does Not Apply	6909
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	46965
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable is designed to provide annual data on required job training for all employed persons. The variable is generated using questions on required formal education and required on-the-job-training which are categorized into up to seven independent variables with 0/1 coding. Out of these, the highest available level of required training is used for the generation of the status variable.
Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
The code (-2) is assigned to all non-employed persons and also includes persons in occupational training, in occupational retraining programs, and those doing an internship at the time of the survey.
Because detailed information onrequired training for job is not assessed in the Questionnaire of the SOEP Innovation Sample, PGAUSB is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGAUSB is coded to "-5" (not contained in questionnaire).

pgpartz – Partner Identicator

0	[0] No Partner in Household	20770
1	[1] Spouse	34974
2	[2] Partner	5929
3	[3] Probably Spouse	76
4	[4] Probably Partner	100
5	[5] unklar	1
-1	[-1] No Answer	401
-2	[-2] Does Not Apply	39
-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

Partner indentifier have the purpose of clearly defining spouse (married) and partner (unmarried) relationships in SOEP-IS households and thus enabling analyses on the couple level. The variable PGPARTZ generated in this context reveals whether a person in a SOEP-IS household has a partner in that household, and if so, the type of relationship existing between the partners. Relationships with persons outside the SOEP household are not covered by this variable.
To explain the codes:
Code 0 is automatically assigned to all persons living in households in which there is clearly no partnership. These include:

(a) one-person households

(b) single-parent households

(c) household head living together with only one parent (or parent-in-law)

Codes 1 to 4 define these relationships. To assign Codes 1 and 2, the partnership has to be clearly definable from the perspective of both partners. This implies agreement between the codes of the variable STELL (= relationship to head of household in PBRUTTO) pointing to a possible partnership (e.g., the combination 0 (=head of household) and 1 (=spouse of household head)), as well as agreement between the codes for family status in that wave (e.g., married couples both have the Code 1 (=married, living together)). In case of ambiguity, the marital history is taken into account as well. If there are inconsistencies between the answers provided by the two persons, or between data on marital status and relationship to head of household, each person is examined individually within his or her household context. If uncertainty remains, the codes 3 or 4 are assigned.

pgpartnr – Person ID number of partner

? Please state the first name of your partner. *(from: soep-is/soep-is-2018-f/Q311:ppnr[])*

-1	[-1] No Answer	0
-2	[-2] Does Not Apply	21210
-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

Partner indicators have the purpose of clearly defining spouse (married) and partner (unmarried) relationships in SOEP households and thus to make possible analyses on the couple level.
If PGPARTZ is coded 0 or 9, this person has no partner or the partner cannot be identified as such. The variable PGPARTNR is assigned the missing code of "-2" (=does not apply) for these persons.
If PGPARTZ is coded 1, 2, 3 or 4, a partnership was defined and PGPARTNR is then assigned the value of the unchanging person ID number (=PERSNR) of the partner.
For analyses of partner relationships, this information can be used to clearly link all persons with their respective partners, and all information on both partners can also be stored in a

common dataset.

pgnation – Citizenship - nationality

? Do you have German citizenship? *(from: soep-is/soep-is-2019/Q179:lsta1a[28193])*
? What is your citizenship? *(from: soep-is/soep-is-2019/Q180:lpnat[28194])*

1	[1] German	57306
2	[2] Turkey	739
3	[3] Ex-Yugoslavia	5
4	[4] Greece	179
5	[5] Italian	308
6	[6] Spain	68
7	[7] Ex-GDR	2
10	[10] Austria	164
11	[11] France	63
12	[12] Belgian, Dutch, Luxembourg	0
13	[13] Denmark	3
14	[14] Great Britain	62
15	[15] Sweden	6
16	[16] Norway	3
17	[17] Finland	13
...	(149 rows omitted)	1523
170	[170] Surinam	0
171	[171] Guyana	0
172	[172] Caucasus	0
173	[173] Zimbabwe	0
174	[174] Madagascar	0
175	[175] Grenada	0
176	[176] Lesotho	0
177	[177] Bhutan	3
196	[196] Kosovo	4
-1	[-1] No Answer	1839
-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

This variable is designed to integrate the information on respondents' nationality for all subsamples.

pgsbil – Diplomas/degrees from secondary/tertiary

? Where did you last attend school? *(from: soep-is/soep-is-2019/Q281:lsab3[28306])*
? What kind of general education / secondary school is it? *(from: soep-is/soep-is-2019/Q293:paus2[28316])*

1	[1] Sec. Gen. School Leaving Certificate	18528
2	[2] Intermediate School Degree	18443
3	[3] Leaving Certificate From Voc High School	3891

4	[4] College Entrance Exam	13560
5	[5] Other	4989
6	[6] Dropout, No School Certificate	1109
7	[7] Currently In School	766
-1	[-1] No Answer	1004
-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

All respondents in all SOEP subsamples are asked about diplomas/degrees attained for completion of secondary/tertiary education the first time they participate in SOEP.
As the SOEP Innovation Sample does not include a youth questionnaire, since 2012 information usually coming from the youth questionnaire was not included in the generation of PGSBIL.

pgbbil01 – Vocational degree attained

? What type of vocational training or university degree did you receive? *(from: soep-is/soep-is-2019/Q289:lab02=2[28312])*

1	[1] Apprenticeship	27994
2	[2] Vocational School	5670
3	[3] Health Care School	197
4	[4] Technical School	3752
5	[5] Civil Servant Training	1578
6	[6] Other Degree	881
7	[7] Completed Vocational Training/Education in Germany	485
-1	[-1] No Answer	978
-2	[-2] Does Not Apply	20755
-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

All respondents in all subsamples are asked about vocational degrees attained the first time they participate in SOEP. The categories that originally each constituted individual variables are combined to make them compatible with the annual question about changes in vocational degrees attained, and this data is updated annually.
Since the SOEP Innovation Sample does not include a youth questionnaire, the information usually coming from the youth questionnaire was not included in the generation of PGBBIL01.

pgbbil02 – Completed college education

? What type of vocational training or university degree did you receive? *(from: soep-is/soep-is-2019/Q289:lab06=1[28312])*

1	[1] Fachhochschule	3562
2	[2] University, Technical College	6573

3	[3] College Not In Germany	148
4	[4] Engineering, Techncial School (East)	868
5	[5] University (East)	761
6	[6] graduation, state doctorate	848
-1	[-1] No Answer	978
-2	[-2] Does Not Apply	48552
-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

All respondents in all subsamples are asked about completed college education the first time they participate in SOEP. To generate the variable, the different degrees/diplomas for all subsamples are integrated. Category (3) "college abroad" is only defined for persons who completed a foreign-language version of the questionnaire. Generation of the variable entails combining the categories to make them compatible with the annual question about changes in vocational degrees/diplomas attained.
Since the SOEP Innovation Sample does not include a youth questionnaire, the information usually coming from the youth questionnaire was not included in the generation of PGBBIL02.

pgbbil03 – No vocational degree

? Did you complete vocational training or a university degree? *(from: soep-is/soep-is-2019/Q288:lab01[28311])*

1	[1] No Vocation Degree	8735
2	[2] Apprenticeship	2755
3	[3] Studies	2524
-1	[-1] No Answer	978
-2	[-2] Does Not Apply	47298
-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

In connection with the question about vocational degrees (PGBBIL01 and PGBBIL02), all first-time respondents to all subsamples are explicitly asked whether they (still) do not possess a vocational degree. In the subsequent years, this data is carried forward or updated. The variable has the Missing Value Code -2 (does not apply) if one of the other two variables on vocational degree has a positive value.
Since the SOEP Innovation Sample does not include a youth questionnaire, the information usually coming from the youth questionnaire was not included in the generation of PGBBIL03.

pgsbila – Secondary school degrees/diplomas abroad

? Where did you last attend school? *(from: soep-is/soep-is-2019/Q281:lsab3[28306])*
? What type of school-leaving certificate did you attain? *(from: soep-is/soep-is-2019/Q283:lsab5[28308])*

1	[1] Mandatory schooling not completed	240
2	[2] Mandatory schooling completed	1679
3	[3] Higher-level secondary school	2954
4	[4] Secondary school completed abroad	0
-1	[-1] No Answer	36
-2	[-2] Does Not Apply	57381
-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

As a supplement to the PGSBIL, this variable provides annually updated data on the highest secondary school degree/diploma attained abroad.
Since the SOEP Innovation Sample does not include a youth questionnaire, the information usually coming from the youth questionnaire was not included in the generation of PGSBILA.

pgbbila – Occupational Training in abroad

1	[1] On-The-Job Training	61
2	[2] Vocational Training	126
3	[3] Vocational School	100
4	[4] College	139
5	[5] Other Training	60
6	[6] Vocational Degree[bbil01] Acquired Abroad	1
7	[7] College Degree Completed Abroad	0
8	[8] Completed Vocational Training/Education Other Country	0
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	14114
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	47689
-6	[-6] Questionnaire Version With Modified Filtering	0

As a supplement to the variable PGBBIL01, this variable gives (and updates) the highest-level vocational degree attained abroad.
Because detailed information on occupational training abroad is not assessed in the Questionnaire of the SOEP Innovation Sample, PGBBILA is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGBBILA is coded to "-5" (not contained in questionnaire).

pgsbilo – Secondary school degree/diploma - East Germany

? Where did you last attend school? *(from: soep-is/soep-is-2019/Q281:lsab3[28306])*

1	[1] 8th Gr. Completed	2538
2	[2] 10th Grade Completed	4741
3	[3] Abitur, EOS	1833
4	[4] Other	120
5	[5] Dropout, No School Certificate	44

6	[6] Currently In School	0
-1	[-1] No Answer	7
-2	[-2] Does Not Apply	53007
-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

As a supplement to the variable PGSBIL the highest secondary school degree/diploma in East Germany is provided as a separate variable.
New SOEP respondents are also asked about secondary degrees/diplomas obtained in the former GDR; and for old respondents, the same codes are carried forward.
As SOEP-IS does not include a youth questionnaire, since 2012 information usually coming from the youth questionnaire was not included in the generation of PGSBILO.

pgbbilo – Vocational degree attained - East Germany

? What type of vocational training or university degree did you receive? *(from: soep-is/soep-is-2019/Q289:lab02=2[28312])*

1	[1] Vocational Training	4775
2	[2] Master Craftsman	589
3	[3] Engineering, Technical Degree	1041
4	[4] Other Degree	101
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	55060
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	724
-6	[-6] Questionnaire Version With Modified Filtering	0

To supplement the variable PGBBIL01 the highest secondary school degree/diploma in East Germany is provided as a separate variable and updated if necessary for 1991. Since 1992 only the West German version has been used for new vocational degrees. For new SOEP respondents, vocational degrees attained in the former GDR are asked as well; for old respondents, the same codes are carried forward. From 2002 on, the questionnaire was expanded and revised, but this led to an operationalization involving more assumptions on the vocational degrees attained in the GDR; (from 2002 on, Code 3 is also listed as the additional category Code 4 in the integrated variables PGBBIL03 if this degree has not been replaced by a more recently attained, higher-level university or college degree).
Since the SOEP-IS does not include a youth questionnaire, the information usually coming from the youth questionnaire was not included in the generation of PGBBILO.

pgfamstd – Marital status in survey year

? What is your marital status? *(from: soep-is/soep-is-2019/Q169:pfamst[28183])*

1	[1] Married	35231
2	[2] Married, But Separated	1275
3	[3] Single	14641
4	[4] Divorced	6496

5	[5] Widowed	4450
6	[6] Registered same sex partnership	109
7	[7] Registered same sex partnership, seperated	17
-1	[-1] No Answer	60
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	11
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

Marital status is describing the institutional status of marriage at the time of the person interview. Marital status is based on information given by the respective person on his or her current relationship as well as on retrospective information about previous relationships asked in the biography questionnaire. For those whose partner was identified within the household, marital status is counter-checked with the information given by the partner. Where contradictions can be found, indication of the person information is compiled if reasonable. If no information is available, the indication by position related to head of household is deferred. Remaining contradictions are solved using information on marriage status when a child was born as well as future reports on a given relationship. Marital status is only available for people, who were interviewed.
Note that the partner indicator PGPARTZ supplied in the PGEN data files as well might not match the information provided in PGFAMSTD in its entirety.

pgbilzt – Amount of education or training (in years)

7		888
8.5		122
9		4888
10		2989
10.5		12872
11		2133
11.5		11191
12		5689
13		3258
13.5		1159
14		1279
14.5		1770
15		3632
16		1613
17		196
18		6835
-1	[-1] No Answer	1006
-2	[-2] Does Not Apply	770
-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 following statements describe the standard computation for amount of education or training (in years).

years of education =years of schooling + years of occupational training
Schooling:

- `no degree` `<= 7 years`
- `lower school degree` `= 9 years`
- `intermediary school` `= 10 years`
- `degree for a professional coll.` `= 12 years`
- `high school degree` `= 13 years`
- `other` `= 10 years`

additional occupational training (includes universities):

- `apprenticeship` `= 1.5 years`
- `technical schools (incl. health)` `= 2 years`
- `civil servants apprenticeship` `= 1.5 years`
- `higher technical college` `= 3 years`
- `university degree` `= 5 years`

Note that for a high school degree 13 years of education are being taken into account, despite the changes of reducing high school by one year in most German federal states in the period 2001-2007.
Furthermore, the introduction of a Bachelor/Master System in the German higher education system in the early 2000´s is not yet reflected in the calculation of years in (higher) education. Hence, 5 years of university education is taken into account although the respondents could have finished in 3 years with the Bachelor's degree.
Helberger, Christof (1988): Eine Überprüfung der Linearitätsannahme der Humankapitaltheorie. In: H.-J. Bodenhöfer (ed.): Bildung, Beruf, Arbeitsmarkt, pp. 151-170, Berlin.
Schwarze, Johannes (1991): Ausbildung und Einkommen von Männern - Einkommensfunktionsschätzungen für die ehemalige DDR und die Bundesrepublik Deutschland. In: Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, (24), pp. 63-69.

pgerwzt – Length Of Time With Firm

? Q208:pstell!=1: Since when have you been working for your current employer? Q208:pstell=1: Since when do you practice your current work? // Month *(from: soep-is/soep-is-2019/Q396:pseitm [28414])*
? Q208:pstell!=1: Since when have you been working for your current employer? Q208:pstell=1: Since when do you practice your current work? // Year *(from: soep-is/soep-is-2019/Q396:pseitj [28414])*

-1	[-1] No Answer	157
-2	[-2] Does Not Apply	27486
-3	[-3] Answer Improbable	67
-4	[-4] Inadmissible Multiple Answer	50
-5	[-5] Not Contained In Questionnaire	48
-6	[-6] Questionnaire Version With Modified Filtering	54

The variable PGERWZT is designed to offer data on the length of time with the firm at the point in time of the interview for all employed persons. This variable is generated from the respondent's start date with the current employer. In the case of a job change within the firm, the full length of time with the firm is calculated. Hence, the variable describes the length of time with the same firm and not the length of time in the same position.
The variable provides consistent longitudinal information on the length of time with the same employer. Data that show longitudinal inconsistencies are corrected.

1. In case of no job change, the information on the start date with the current employer given in the earliest interview available is treated as dominant and carried forward to the subsequent years.

2. In case of a job change between firms, the information on the start of the current position is used and carried forward to the subsequent years.

3. Up to wave Z (2009), a respondent who starts working again after a period of non-employment is assumed to have returned to the former employer if the indicated start date with the current employer was before the previous interview date. In this case, the start date with the current employer given in the previous interview is treated as dominant. Otherwise, the present information on the start date with the current employer is used and carried forward to the subsequent years. For respondents who are assumed to have returned to their former employer, the full length of time with the firm is calculated. There is no deduction for the time during which the respondent was not employed.

4. Since wave BA (2010), there is a modified answer category in the questionnaire which indicates that a respondent returns to his/her former employer after a period of non-employment. If a respondent indicates to have started working again at a former employer, the present information on the start date with the current employer is used and carried forward to the subsequent years. Unlike before wave BA (2010), the present information is treated as dominant even if the indicated start date with the current employer was before the previous interview date. Hence, the full length of time with the firm is calculated, and there is no deduction for the time during which the respondent was not employed or employed in another firm.

5. The length of time with the firm is also provided for the East German sample since its start in 1990. Due to the massive restructuring of the economy that took place in East Germany after reunification, this variable should be dealt with cautiously in the first transition years.

Both monthly and annual information is used in the variables and rounded off as length of time in years (with months in decimal form). If the month was not available a random month is used.

pgtatzt – Actual Weekly Work Time

? And how many hours do you generally work per week, including any overtime? // [Whole number] Hours per week *(from: soep-is/soep-is-2019/Q431:paz10a[28447])*

0.1	3
0.2	5
0.3	12
0.4	15

0.5		18
0.6		14
0.7		15
0.8		21
0.9		10
1		100
1.1		11
1.2		39
1.3		5
1.4		5
1.5		38
...	(209 rows omitted)	32085
70		196
71		1
72		21
73		3
75		37
76		4
77		1
78		2
80		14
-1	[-1] No Answer	1450
-2	[-2] Does Not Apply	26421
-3	[-3] Answer Improbable	100
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	1644
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable is designed to offer annual data on actual weekly working hours (including overtime) for all persons employed at the time of the survey (including the self-employed). The data are obtained by asking respondents how many hours they work on average per week.
For implausible answers (actual weekly working hours of more than 80 per week), we assign the value (-3). The variable is rounded off and gives the number of working hours as a decimal number.
Please also see PGVEBZT and PGUEBSTD.

pgvebzt – Agreed Upon Weekly Work Time

? How many hours per week are stipulated in your contract (excluding overtime)? // [Whole number] Hours per week *(from: soep-is/soep-is-2019/Q430:paz08a[28446])*

0.3		5
0.4		8
0.5		13
0.6		13
0.7		10
0.8		19
0.9		11
1		59
1.1		10

1.2		31
1.3		3
1.4		2
1.5		41
1.6		22
1.7		3
...	(219 rows omitted)	27526
57		1
57.5		1
60		36
62.5		1
65		5
70		3
72		1
75		2
78		1
-1	[-1] No Answer	1869
-2	[-2] Does Not Apply	30947
-3	[-3] Answer Improbable	3
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	1644
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable is designed to offer annual data on agreed weekly working hours. The variable takes into account only those persons who were in dependent employment (not self-employed) at the time of the survey. Agreed weekly working hours were asked up to 1989 only in full hours, and from 1990 on in three-digit form (counting the first digit after the decimal point).
The value (-2) is assigned to non-employed people, employees without set hours and to self-employed people, including self-employed farmers, freelancers, and other self-employed persons. In 2012, the value (-2) was assigned only to non-employed people and to self-employed people, including self-employed farmers, freelancers, and other self-employed persons. If persons helping out in family businesses report agreed weekly working hours, we assign a non-missing value.
For implausible answers (agreed weekly working time of more than 80 hours per week) we assign the value (-3).
The variable is rounded off and gives the number of working hours as a decimal number.
Please also see PGTATZT and PGUEBSTD.

pguebstd – Overtime per Week

0		14246
0.1		89
0.2		279
0.3		159
0.4		126
0.5		348
0.6		60
0.7		123
0.8		42

0.9		145
1		683
1.1		12
1.2		221
1.3		19
1.4		141
...	(130 rows omitted)	11448
30.5		1
31		1
32		1
32.5		1
33		3
34		2
35		8
40		1
44		1
-1	[-1] No Answer	2527
-2	[-2] Does Not Apply	29957
-3	[-3] Answer Improbable	2
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	1644
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable is designed to offer annual data on overtime per week for all persons in dependent employment at the time of the survey. The data is obtained by asking respondents how many overtime hours they worked in the month before the survey. The number of monthly overtime hours is then converted into weekly overtime by dividing the number given by 4.3. Since PGUEBSTD refers to weekly overtime during the last month, the number may deviate from the difference between average actual weekly working hours and the agreed weekly working hours.
In the year 2012, respondents were not asked about the number of hours of overtime per week. PGUEBSTD was therefore generated using the difference between average actual weekly working hours and agreed weekly working hours.
The value (-2) is assigned to non-employed people, employees without set hours and to self-employed people, including self-employed farmers, freelancers, and other self-employed persons. If persons helping out in family businesses report overtime hours, we assign a non-missing value. For implausible answers (agreed-upon weekly working time or actual weekly working time of more than 80 hours per week AND weekly overtime of more than 10 hours we assign the value (-3).
The variable is rounded down and gives the number of overtime hours as a decimal.
Please also see PGVEBZT and PGTATZT.

pglfs – Labor Force Status

? Have you been engaged in paid work during the last 7 days, even if this work was only for an hour or just a few hours? *(from: soep-is/soep-is-2019/Q269:p7tag[28297])*

? IF [Woman & age <= 49] Are you currently on maternity leave or legislatively regulated parental leave (“Elternzeit”)? IF [Man OR (Woman & age > 49)] Are you currently on legislatively regulated parental leave (“Elternzeit”)? *(from: soep-is/soep-is-2019/Q270:perz[28298])*

? Are you currently enrolled in an educational or training program? In other words: are you in school or higher education, working on a doctors degree, completing vocational training, or taking part in further training? *(from: soep-is/soep-is-2019/Q291:paus1[28314])*
? Are you currently employed? Which one of the following applies best to your status? *(from: soep-is/soep-is-2019/Q302:perw[28325])*

1	[1] Non-Working	5727
2	[2] NW-Age 65 And Older	13678
3	[3] NW-In Education-Training	1479
4	[4] NW-Maternity Leave	826
5	[5] NW-Military-Community Service	30
6	[6] NW-Unemployed	2702
8	[8] NW-But Sometimes Sec. Job	216
9	[9] NW-work but past 7 days	523
10	[10] NW-But Reg. Sec. Job	2024
11	[11] Working	34611
12	[12] Working But NW Past 7 Days	473
-1	[-1] No Answer	1
-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

This variable is based on the annual question on current employment status, combined with additional information on activities of non-working individuals. Since the beginning of the SOEP in the year 1984, the number of values assigned has been based on a large number of highly differentiated answer categories. It is designed to provide consistent longitudinal data on labor force participation across all waves.
PGLFS provides a differentiation between “working” (Code 11-12) and “non-working” (Code 1-10), categories which are constant over all waves. Non-employment is subdivided further in order to make it possible to efficiently apply different labor market concepts in studying the data. To calculate this variable, the variables on employment status, age, maternity leave, second jobs, registration at the employment office, participation in paid work during the past 7 days and training status are used. Code (12) was added in 2000.
For respondents who have multiple status codes and different values for this variable, the following hierarchy was used to determine which of the values would play the determining role (increasing dominance):

```
11 - working

 1 - non-working without further information

 2 - non-working, and older than 65

 3 - non-working, and currently in a training program

 6 - non-working, and registered unemployed

 4 - non-working, on maternity leave
```

```
  5 - non-working, in military/community service

  9 - non-working, but working past 7 days

 10 - non-working, but regular second job

  8 - non-working, but occasional second job

 12 - working, but non-working past 7 days
```

PGLFS supplements the variable PGEMPLST, which differentiates among persons who are employed.

pgis88 – Current Occupational Classification (ISCO-88 Com)

0	[0] Soldiers	0
100	[100] Soldiers	78
1000	[1000] Legislators, Senior Officials and Managers	0
1100	[1100] Legislators and Senior Government Officials	0
1110	[1110] Legislators and Senior Government Officials	1
1140	[1140] Senior Officials of Special-Interest Organisations	5
1141	[1141] Senior Officials of Political Party Organisations	0
1142	[1142] Senior Officials of Employers', Workers' and Other Economic-Interest Organisations	26
1143	[1143] Senior Officials of Humanitarian and Other Special-Interest Organisations	0
1200	[1200] Corporate Managers	84
1210	[1210] Directors and Chief Executives	233
1220	[1220] Production and Operations Managers	0
1221	[1221] Production and Operations Managers in Agriculture, Hunting, Forestry and Fishing	0
1222	[1222] Production and Operations Managers in Manufacturing	85
1223	[1223] Production and Operations Managers in Construction	0
...	(469 rows omitted)	23787
9213	[9213] Fishery, Hunting and Trapping Labourers	0
9300	[9300] Labourers in Mining, Construction, Manufacturing and Transport	0
9310	[9310] Mining and Construction Labourers	0
9311	[9311] Mining and Quarrying Labourers	2
9312	[9312] Construction and Maintenance Labourers: Roads, Dams and Similar Constructions	3
9313	[9313] Building Construction Laborer	39
9320	[9320] Manufacturing Laborer	366
9330	[9330] Transport Lab., Freight Handler	178
-1	[-1] No Answer	308
-2	[-2] Does Not Apply	20300
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	16795
-6	[-6] Questionnaire Version With Modified Filtering	0
-8	[-8] Question this year not part of Survey program	0

Code name (Main group, group):

(1000) Legislators, senior officials, and managers

(1001) Legislators and senior officials

(1002) Corporate managers

(1003) Managers of small enterprises

(1004) Professionals

(1005) Physical, mathematical, and engineering science professionals

(1006) Life science and health professionals

(1007) Teaching professionals

(1008) Other professionals

(1009) Technicians and associate professionals

(1010) Physical and engineering science associate professionals

(1011) Life science and health associate professionals

(1012) Teaching associate professionals

(1013) Other associate professionals

(1014) Clerks

(1015) Office clerks

(1016) Customer services clerks

(1017) Service Workers and shop and market sales workers

(1018) Personal and protective services workers

(1019) Models, salespersons, and demonstrators

(1020) Skilled agricultural and fishery Workers

(1021) Skilled agricultural and fishery workers

(1022) Craft and related trades workers

(1023) Extraction and building trades workers

(1024) Metal, machinery, and related trades workers

(1025) Precision, handicraft, craft printing and related trades workers

(1026) Other craft and related trades workers

(1027) Plant and machine operators and assemblers

(1028) Stationary plant and related operators

(1029) Machine operators and assemblers

(1030) Drivers and mobile plant operators

(1031) Elementary occupations

(1032) Sales and services elementary occupations

(1033) Agricultural, fishery, and related laborers

(1034) Laborers in mining, construction, manufacturing, and transport

This variable is designed to provide annual data on occupational activity for all employed persons according to the International Standard Classification of Occupations ISCO-88. Respondents answer the question on their current occupational title in their own words, and this response is entered into a blank in the questionnaire.
ISCO-88 is a strictly four-digit classification, and this variable is therefore coded in four-digit form. In contrast to the previous version of the classification system, ISCO-68, ISCO-88 does not use blanks if there is no adequate information for specific coding, but uses zeros instead. Thus 4000 stands for an unspecified office job; 2300 stands for teachers and 2000 stands for scientists, both without closer specification. There is no conversion key since the two classifications differ significantly. Hartmann and Schütz (2002) provide detailed information on the conducted occupational coding. This result has been slightly modified to fit to the ISCO-88 version for European Union purposes (ISCO-88(COM)).
Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
In the SOEP-IS, ISCO-88 (pgis88) is released up until 2016 and coded to [-5] afterwards. Starting in 2017, only ISCO-08 scores will be released (see pgis08).
Hartmann/Schütz (2002): Die Klassifikation der Berufe und der Wirtschaftszweige im Sozioökonomischen Panel. Neuvercodung der Daten 1984–2001. Infratest Sozialforschung, München. https://www.diw.de/documents/dokumentenarchiv/17/diw_01.c.40132.de/vercodung.pdf

pgis08 – Current Occupational Classification (ISCO-08)

0	[0] Armed Forces Occupations	0
100	[100] Commissioned Armed Forces Officers	0
110	[110] Commissioned Armed Forces Officers	4
200	[200] Non-commissioned Armed Forces Officers	0
210	[210] Non-commissioned Armed Forces Officers	2
300	[300] Armed Forces Occupations, Other Ranks	0
310	[310] Armed Forces Occupations, Other Ranks	9
1000	[1000] Managers	0
1100	[1100] Chief Executives, Senior Officials and Legislators	0
1110	[1110] Legislators and Senior Officials	0
1111	[1111] Legislators	2
1112	[1112] Senior Government Officials	19
1113	[1113] Traditional Chiefs and Heads of Villages	0
1114	[1114] Senior Officials of Special-interest Organizations	3
1120	[1120] Managing Directors and Chief Executives	109
...	(566 rows omitted)	12497
9611	[9611] Garbage and Recycling Collectors	15
9612	[9612] Refuse Sorters	5
9613	[9613] Sweepers and Related Labourers	1
9620	[9620] Other Elementary Workers	0
9621	[9621] Messengers, Package Deliverers and Luggage Porters	53
9622	[9622] Odd-job Persons	0

9623	[9623] Meter Readers and Vending-machine Collectors	1
9624	[9624] Water and Firewood Collectors	0
9629	[9629] Elementary Workers Not Elsewhere Classified	22
-1	[-1] No Answer	851
-2	[-2] Does Not Apply	10299
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	38398
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable is designed to provide annual data on occupational activity for all employed persons according to the International Standard Classification of Occupations ISCO-08 (Version of the ILO). Respondents answer the question on their current occupational title in their own words, and this response is entered into a blank in the questionnaire. ISCO-08 is a strictly four-digit classification, and this variable is therefore coded in four-digit form. Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
ISCO-08 was adopted through a resolution of a Tripartite Meeting of Experts on Labour Statistics held in December 2007. This resolution was subsequently endorsed by the Governing Body of the ILO in March 2008.
Sourcelink: http://www.ilo.org/public/english/bureau/stat/isco/isco08/index.htm
Source of German Labels http://ec.europa.eu/eurostat/ramon/documents/SCL/isco08/SCL_isco08.zip
In the SOEP-IS, ISCO-08 (pgis08) was first implemented in wave 2016 and is coded [-5] in all previous waves.

pgisei88 – ISEI-Status following Ganzeboom (based on IS88)

-1	[-1] No Answer	294
-2	[-2] Does Not Apply	17231
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	16795
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable reflects the Standard International Socio-Economic Index of Occupational Status for all employed persons. The ISEI Index was developed in 1992 by Ganzeboom, De Graaf, Treiman, and De Leew based on information about income, education, and occupation. Technically, ISEI was created by scaling the ISCO-88 classification. The values for the variable range between 16 and 90. In contrast to the prestige scores of Ganzeboom and Treiman (1996) and Wegener (1988), ISEI is a measure of socio-economic status. It is derived from the ISCO-88 code of the current occupation using the Stata ado iskoisei by John Hendrick which itself is based on Harry Gaanzeboom's SPSS algorithms. Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
Please also see occupational prestige scores (pgsiops88, pgmps92) and occupational class (pgegp88).
In the SOEP-IS, pgisei88 is released up until 2016 and coded to [-5] afterwards. Starting in 2017, only pgisei08 scores will be released.
Ganzeboom, H. B. G. / De Graaf, P .M. / Treiman, D. J. / De Leew, J.(1992): A Standard Inter-

national Socio-Economic Index of Occupation Status, In: Social Science Research 21: 1-56

pgisei08 – ISEI-Status following Ganzeboom (based on IS08)

-1	[-1] No Answer	821
-2	[-2] Does Not Apply	5733
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	42964
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable reflects the Standard International Socio-Economic Index of Occupational Status for all employed persons. The ISEI Index was developed in 1992 by Ganzeboom, De Graaf, Treiman, and De Leew based on information about income, education, and occupation. Technically, ISEI was created by scaling the ISCO-88 classification. The values for the variable range between 11 and 90. In contrast to the prestige scores of Ganzeboom and Treiman (1996) and Wegener (1988), ISEI is a measure of socio-economic status. Starting in SOEP-IS wave 2016, pgisei08 is derived from ISCO-88 scores that themselves are derived from ISCO-08 scores using the „derivescores“ STATA package (https://github.com/dirtyhawk/stata-derivescores).
This recoding is derived from Ganzeboom‘s SPSS script iskoisei.sps. (http://www.harryganzeboom.nl/isco88/in
Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
Please also see occupational prestige scores (pgsiops08, pgmps08) and occupational class (pgegp08).
In the SOEP-IS, pgisei08 was first implemented in wave 2016 and is coded [-5] in all previous waves.
Ganzeboom, H. B. G. / De Graaf, P .M. / Treiman, D. J. / De Leew, J.(1992): A Standard International Socio-Economic Index of Occupation Status, In: Social Science Research 21: 1-56

pgmps92 – Magnitude Prestige Scale (based on KldB92)

30		11
30.1		34
30.2		54
30.3		5
31		261
31.1		72
31.2		7
31.5		266
31.7		141
31.8		17
31.9		4
32		24
32.1		277
32.2		14
32.3		808
...	(157 rows omitted)	24968
135.7		194

138.2		34
138.9		14
139.8		45
145.7		149
152.5		170
191.3		294
207.2		41
216		17
-1	[-1] No Answer	335
-2	[-2] Does Not Apply	17239
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	16795
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable gives the occupational prestige score developed by Wegener (1988) for all employed persons. Like the PGSIOPS prestige sore, Wegener's prestige scala measures a person's occupational prestige and was developed especially for use in the Federal Republic of Germany. PGMPS is assigned based on the German Federal Statistical Office's occupational classification of 1992 (PGKLAS). The procedure has been documented in Frietsch and Wirth (2001). Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
Please also see occupational prestige scores (pgsiops88), occupational status (pgisei88), and occupational class (pgegp88).
In the SOEP-IS, pgmps92 is released up until 2016 and coded to [-5] afterwards. Starting in 2017, only pgmps08 scores will be released.
Wegener, Bernd (1988): Kritik des Prestiges, Opladen.
Frietsch, Rainer, and Heike Wirth (2001): Die Übertragung der Magnitude-Prestigeskala von Wegener auf die Klassifikation der Berufe. In: ZUMA Nachrichten 48 (Jg.25): 139-165.

pgmps08 – Magnitude Prestige Scale (based on IS88 recoded from IS08)

20		34
23.9		29
24.7		15
26.7		29
26.9		133
28.6		1
30		390
30.3		20
31.2		59
31.6		6
31.8		84
31.9		33
32.4		60
34.7		3
35.6		8
...	(155 rows omitted)	10548
153.5		13

153.8		108
159.8		128
160.3		21
160.5		49
170.9		42
173.3		21
179.6		139
186.8		7
-1	[-1] No Answer	1613
-2	[-2] Does Not Apply	5733
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	42964
-6	[-6] Questionnaire Version With Modified Filtering	0

The Magnitude-Prestige-Scale (MPS) is a tool for operationalising the social position of occupations in empirical surveys. In contrast to alternative instruments such as the Standard International Occupational Prestige Scale (SIOPS) or the International Socio-economic Index of Occupational Status (ISEI), the MPS is an instrument that was constructed for analyses on the national level only (Christoph, 2005). Starting in SOEP-IS wave 2016, pgmps08 is derived from ISCO-88 scores that themselves are derived from ISCO-08 scores using the „derivescores“ STATA package (https://github.com/dirtyhawk/stata-derivescores). Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
Please also see occupational prestige scores (pgsiops08), occupational status (pgisei08), and occupational class (pgegp08).
In the SOEP-IS, pgmps08 was first implemented in wave 2016 and is coded [-5] in all previous waves.

pgnace – Two-digit NACE Industry-Sector (NACE Rev. 1.1, Sector)

? In which branch of business or industry is your company or institution active for the most part? *(from: soep-is/soep-is-2019/Q315:pbra[28335])*

1	[1] Agriculture, Hunting, Related Service Activities	517
2	[2] Forestry, Logging, Related Service activities	26
5	[5] Fishing, Operation Of Fish Hatcheries And Fish Farms	8
10	[10] Mining Of Coal And Lignite; Extraction Of Peat	89
11	[11] Extraction Of Crude Petroleum And Natural Gas	1
12	[12] Mining Of Uranium And Thorium Ores	0
13	[13] Mining Of Metal Ores	0
14	[14] Other Mining And Quarrying	19
15	[15] Manuf Food Products And Beverages	539
16	[16] Manuf Tobacco Products	9
17	[17] Manuf Textiles	101
18	[18] Manuf Wearing Apparel; Dressing And Dyeing Of Fur	59
19	[19] Tanning,Dressing Of Leather; Manuf luggage, Footwear	10
20	[20] Manuf Wood Products, Except Furniture	93
21	[21] Manuf Pulp, Paper And Paper Products	81
...	(40 rows omitted)	24724

91	[91] Activities Of Membership Organizations NEC.	424
92	[92] Recreational, Cultural And Sporting Activities	626
93	[93] Other Service Activities	274
95	[95] Private Households With Employed Persons	132
96	[96] Industry - NEC	144
97	[97] Handcraft, Trade - NEC	128
98	[98] Services - NEC	296
99	[99] Extra-territorial Organizations And Bodies	10
100	[100] Manufacturing - NEC	79
-1	[-1] No Answer	962
-2	[-2] Does Not Apply	22322
-3	[-3] Answer Improbable	1
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	10616
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable is designed to provide annual data on the industry of economic activity for all employed persons according to the Statistical Classification of Economic Activities in the European Community (Nomenclature des statistiques des activités économiques de la Communauté européenne - NACE). Respondents answer the question in their own words regarding the industry in which they are currently working, and this response is entered into a blank in the questionnaire. In order to facilitate international comparability, the European industry standard classification system is used by Infratest Sozialforschung to recode this information. This recoding has been documented in Hartmann/Schütz 2002.
The codes in NACE Rev.1 also correspond to ISIC Rev.3 (International Standard Classification of All Economic Activities). With the 2001 data distribution, the sector codes formerly used in the SOEP were completely recoded to the NACE classification. Please note that special codes 96-98 as well as 100 were assigned by Infratest Sozialforschung whenever respondents did not provide a more detailed answer.
Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
Hartmann/Schütz (2002): Die Klassifikation der Berufe und der Wirtschaftzweige im Sozio-oekonomischen Panel – Neuvercodung der Daten 1984 – 2001. Infratest Sozialforschung, München.

pgnace2 – Two-digit NACE Industry-Sector (NACE Rev. 2, Sector)

1	[1] Crop and animal production, hunting and related service activities	65
2	[2] Forestry and logging	6
3	[3] Fishing and aquaculture	1
5	[5] Mining of coal and lignite	9
6	[6] Extraction Of Crude Petroleum And Natural Gas	3
7	[7] Mining Of Metal Ores	0
8	[8] Other Mining And Quarrying	2
9	[9] Mining support service activities	0
10	[10] Manufacture of food products	98
11	[11] Manufacture of beverages	10
12	[12] Manuf Tobacco Products	4
13	[13] Manuf Textiles	6

14	[14] Manuf Wearing Apparel; Dressing And Dyeing Of Fur	18
15	[15] Manufacture of leather and related products	3
16	[16] Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	13
...	(65 rows omitted)	5448
92	[92] Gambling and betting activities	3
93	[93] Sports activities and amusement and recreation activities	23
94	[94] Activities of membership organisations	99
95	[95] Repair of computers and personal and household goods	10
96	[96] Other personal service activities	51
97	[97] Private Households With Employed Persons	33
98	[98] Undifferentiated goods- and services-producing activities of private households for own use	0
99	[99] Extra-territorial Organizations And Bodies	2
-1	[-1] No Answer	48
-2	[-2] Does not apply	4661
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	51674
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

pgsiops88 – TREIMANS STANDARD INT.OCC.PR.SCORE (based on IS88)

13		17
15		57
17		5
18		2
19		474
20		351
21		1129
22		175
23		275
24		51
25		442
26		32
27		22
28		220
29		160
...	(37 rows omitted)	23331
68		7
69		30
70		411
71		58
72		81
73		116
75		19
76		30
78		475
-1	[-1] No Answer	294
-2	[-2] Does Not Apply	17231

-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	16795
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable gives the occupational prestige score index for all employed persons. PGSIOPS is based on ISCO-88 and was developed by Donald Treiman et al. The scale ranges from 6 to 78.
Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
Please also see occupational prestige scores (pgmps92), occupational status (pgisei88), and occupational class (pgegp88).
In the SOEP-IS, pgsiops88 is released up until 2016 and coded to [-5] afterwards. Starting in 2017, only pgsiops08 scores will be released.
Ganzeboom, Harry B.G. and Donald Treiman (1996): Internationally comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations. In: Social Science Research, Vol. 25, 201-239

pgsiops08 – TREIMANS STANDARD INT.OCC.PR.SCORE (based on IS08)

13		21
15		15
16		33
20		374
20.03		13
20.3		79
20.39		54
21		22
21.08		112
21.67		166
22		83
22.28		7
22.69		9
22.9		15
23		74
...	(171 rows omitted)	11215
70		50
72		39
73.1		42
73.51		60
75		10
75.68		5
76.11		7
78.01		139
78.16		128
-1	[-1] No Answer	821
-2	[-2] Does Not Apply	5733
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0

-5	[-5] Not Contained In Questionnaire	42964
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable gives the occupational prestige score index for all employed persons. PGSIOPS is based on ISCO-88 and was developed by Donald Treiman et al. The scale ranges from 6 to 78. Starting in SOEP-IS wave 2016, pgsiops08 is derived from ISCO-88 scores that themselves are derived from ISCO-08 scores using the „derivescores" STATA package (https://github.com/dirtyhawk/stata-derivescores). This recoding is derived from Ganzeboom's SPSS script iskotrei.sps. (http://www.harryganzeboom.nl/isco88/index.htm). Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
Please also see occupational prestige scores (pgmps08), occupational status (pgisei08), and occupational class (pgegp08).
In the SOEP-IS, pgsiops08 was first implemented in wave 2016 and is coded [-5] in all previous waves.
Ganzeboom, Harry B.G. and Donald Treiman (1996): Internationally comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations. In: Social Science Research, Vol. 25, 201-239

pgegp88 – ERIKSON and GOLDTHORPE Class Category (based on IS88)

1	[1] [I] Higher Managerial and Professional Workers	3873
2	[2] [II] Lower Managerial and Professional Workers	6853
3	[3] [IIIa] Routine Clerical Work	3567
4	[4] [IIIb] Routine Service and Sales Work	3818
5	[5] [IVa] Small Self-Employed With Employees	470
6	[6] [IVb] Small Self-Employed Without Employees	762
7	[7] [V] Manual Supervisors	0
8	[8] [VI] Skilled Manual Workers	4070
9	[9] [VIIa] Semi- and Unskilled Manual Workers	3923
10	[10] [VIIb] Agricultural Labour	374
11	[11] [IVc] Self-Employed Farmers	155
-1	[-1] No Answer	292
-2	[-2] Does Not Apply	17338
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	16795
-6	[-6] Questionnaire Version With Modified Filtering	0
-8	[-8] Question this year not part of Survey program	0

This variable gives the occupational class for all employed persons. PGEGP is derived from the Standard International Socio-Economic Index of Occupational Status (ISEI). Technically, the variable was created by scaling the ISCO-88 classification. In addition, it is based on information about income, education and occupation. The EGP Index was documented by Ganzeboom/Treiman in 1996 and revised in 2003.
Former versions and waves contained additional categories for unemployed persons (15) and pensioners (18). From wave 2015 on the egp-variable has a more standard shape. Information on unemployment and retirement can be found in PGSTIB (occupational position) and PGLFS (labor force status).

Annual information on the occupational position is used to generate the EGP-categories for the self-employed. In case no information on the number of employees is available, the PGEGP-categories (5) and (6) contain information on the firm size for self-employed persons.
Based on the new classification developed by Ganzeboom/Treiman (2003), several ISCO values were recoded in PGEGP as follows:

- ISCO 2470 becomes EGP=1.
- ISCO 2500 becomes EGP=2.
- ISCO 4300, 4400, 4500 become EGP=4.
- ISCO 7900 becomes EGP=7.
- ISCO 9910-9990 become EGP=9.

Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
Please also see occupational status (pgisei88) and occupational prestige scores (pgsiops88, pgmps92).
In the SOEP-IS, pgegp88 is released up until 2016 and coded to [-5] afterwards. Starting in 2017, only pgegp08 scores will be released.
Ganzeboom, H. B. G. /Treiman, D. J. (1996): Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations In: Social Science Research 25: 201-239
Ganzeboom, H. B. G. /Treiman, D. J. (2003): Three Internationally Standardised Measures for Comparative Research on Occupational Status. In: Hoffmeyer-Zlotnik, J. H. P. Wolf, C. (eds.): Advances in Cross-National Comparison. A European Working Book for Demographic and Socio-Economic Variables. New York: Kluwer Academic/ Plenum Publishers. pp. 159–193.

pgegp08 – ERIKSON and GOLDTHORPE Class Category (based on IS88 recoded from IS08)

1	[1] [I] Higher Controllers	1885
2	[2] [II] Lower Controllers	3476
3	[3] [IIIa] Routine Nonmanual	2168
4	[4] [IIIb] Lower Sales-Service	1487
5	[5] [IVa] Selfempl with empl	188
6	[6] [IVb] Selfempl no empl	306
7	[7] [V] Manual Supervisors	37
8	[8] [VI] Skilled Worker	1189
9	[9] [VIIa] Unskilled Worker	1844
10	[10] [VIIb] Farm Labor	130
11	[11] [IVc] Selfempl Farmer	62
-1	[-1] No Answer	821
-2	[-2] Does Not Apply	5733
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	42964
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable gives the occupational class for all employed persons. PGEGP is derived from the Standard International Socio-Economic Index of Occupational Status (ISEI). Technically, the variable was created by scaling the ISCO-88 classification. In addition, it is based on information about income, education and occupation. The EGP Index was documented by Ganzeboom/Treiman in 1996 and revised in 2003. Starting in SOEP-IS wave 2016, pgegp08 is derived from ISCO-88 scores that themselves are derived from ISCO-08 scores using the „derivescores" STATA package (https://github.com/dirtyhawk/stata-derivescores). This version of Goldthorpe's class categories (EGP) is derived from Ganzeboom's SPSS script isko-egp.sps.
Former versions and waves contained additional categories for unemployed persons (15) and pensioners (18). From wave 2015 on the egp-variable has a more standard shape. Information on unemployment and retirement can be found in PGSTIB (occupational position) and PGLFS (labor force status).
Annual information on the occupational position is used to generate the EGP-categories for the self-employed. In case no information on the number of employees is available, the PGEGP-categories (5) and (6) contain information on the firm size for self-employed persons.
Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
Please also see occupational status (pgisei08) and occupational prestige scores (pgsiops08, pgmps08).
In the SOEP-IS, pgegp08 was first implemented in wave 2016 and is coded [-5] in all previous waves.
Ganzeboom, H. B. G. /Treiman, D. J. (1996): Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations In: Social Science Research 25: 201-239
Ganzeboom, H. B. G. /Treiman, D. J. (2003): Three Internationally Standardised Measures for Comparative Research on Occupational Status. In: Hoffmeyer-Zlotnik, J. H. P. Wolf, C. (eds.): Advances in Cross-National Comparison. A European Working Book for Demographic and Socio-Economic Variables. New York: Kluwer Academic/ Plenum Publishers. pp. 159–193.

pgklas92 – Current Occupational Classification (KldB92)

110	[110] Farmers, general	128
111	[111] Fruit and vegetable farmers (non-horticultural)	4
112	[112] Arable farmers (special, permanent crops)	0
113	[113] Livestock farmers and pasture farmers	0
114	[114] Seed, crop producers, propagators (non-horticultural)	0
115	[115] Crop protectors	3
116	[116] Farmers and landlords	0
118	[118] Farmers and wine growers	0
120	[120] Wine growers, general	0
121	[121] Vine propagators	0
129	[129] Other wine growers	0
130	[130] Agricultural workers, general	12
131	[131] Agricultural supervisors	0
132	[132] Agricultural machinery drivers	8
133	[133] Vineyard workers	0
...	(2262 rows omitted)	24689

9832	[9832] Other employees (job-searching) with (as yet) undefined occupation	0
9911	[9911] Specialized professionals without further specification	6
9921	[9921] Homeworkers without further specification	2
9931	[9931] Forepersons, group leaders without further specification	15
9941	[9941] Persons doing community service without further specification	0
9951	[9951] Self-employed persons without further specification	36
9961	[9961] Consultancy, planning professionals without further specification	18
9971	[9971] Other employees without further specification	124
-1	[-1] No Answer	159
-2	[-2] Does not apply	20291
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	16795
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

This variable is designed to provide annual data on job classification for all employed persons according to the classification of the German Federal Statistical Office (StaBuA). Respondents answer the question on their current occupational title in their own words, and this response is entered into a blank in the questionnaire. Due to data protection regulations, this information cannot be provided to data users and was therefore completely recoded by Infratest Sozialforschung in the year 2002. This recoding has been documented in Hartmann/Schütz 2002.
Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
In the SOEP-IS, pgklas92 is released up until 2016 and coded to [-5] afterwards. Starting in 2017, only pgklas10 scores will be released.
The occupational classification of the German Federal Statistical Office differentiates among six main occupational types (see next page):
I KLAS-Codes 0100-0629 Berufe in der Land-, Tier-, Forstwirtschaft und im Gartenbau
II KLAS-Codes 0700-0809 Bergleute, Mineralgewinner
III Fertigungsberufe
IIIa KLAS-Codes 1000-1129 Berufe in der Steinbearbeitung und Baustoffherstellung
IIIb KLAS-Codes 1200-1359 Keramik-, Glasberufe
IIIc KLAS-Codes 1400-1539 Chemie-, Kunststoffberufe
IIId KLAS-Codes 1600-1799 Berufe in der Papierherstellung, -verarbeitung und im Druck
IIIe KLAS-Codes 1800-1859 Berufe in der Holzverarbeitung, Holz- und Flechtwarenherstellung
IIIf KLAS-Codes 1900-2459 Berufe in der Metallerzeugung und –bearbeitung
IIIg KLAS-Codes 2500-3099 Metall-, Maschinenbau- und verwandte Berufe
IIIh KLAS-Codes 3100-3189 Elektroberufe
IIIi KLAS-Codes 3200-3239 MontiererInnen und Metallberufe, a.n.g.
IIIk KLAS-Codes 3300-3619 Textil- und Bekleidungsberufe
IIIl KLAS-Codes 3700-3789 Berufe in der Lederherstellung, Leder- und Fellverarbeitung
IIIm KLAS-Codes 3900-4359 Ernährungsberufe
IIIn KLAS-Codes 4400-4729 Hoch-, Tiefbauberufe
IIIo KLAS-Codes 4800-4929 Ausbauberufe, PolsterInnen
IIIp KLAS-Codes 5000-5069 Berufe in der Holz- und Kunststoffverarbeitung
IIIq KLAS-Codes 5100-5149 MalerInnen, LackiererInnen und verwandte Berufe
IIIr KLAS-Codes 5200-5239 WarenprüferInnen, VersandfertigmacherInnen

IIIs KLAS-Codes 5300-5319 HilfsarbeiterInnen ohne nähere Tätigkeitsangabe
IIIt KLAS-Codes 5400-5509 MaschinistInnen und zugehörige Berufe
IV Technische Berufe
IVa KLAS-Codes 6000-6129 IngenieurInnen, ChemikerInnen, PhysikerInnen, MathematikerInnen
IVb KLAS-Codes 6200-6529 TechnikerInnen, Technische Sonderfachkräfte
V Dienstleistungsberufe
Va KLAS-Codes 6600-6899 Warenkaufleute
Vb KLAS-Codes 6900-7069 Dienstleistungskaufleute und zugehörige Berufe
Vc KLAS-Codes 7100-7449 Verkehrsberufe
Vd KLAS-Codes 7500-7899 Organisations-, Verwaltungs-, Büroberufe
Ve KLAS-Codes 7900-8149 Ordnungs- und Sicherheitsberufe
Vf KLAS-Codes 8200-8399 Schriftwerkschaffende, -ordnende und künstlerische Berufe
Vg KLAS-Codes 8400-8599 Gesundheitsdienstberufe
Vh KLAS-Codes 8600-8949 Sozial- und Erziehungsberufe, anderweitig nicht genannte geistes- und sozialwissenschaftliche Berufe
Vi KLAS-Codes 9000-9379 Sonstige Dienstleistungsberufe
VI KLAS-Codes 9700-9979 Sonstige Arbeitskräfte
Because of gaps in the answers provided by respondents, the following "new" codes were created:
9711 - Mithelfende Familienangehörige außerhalb der Landwirtschaft, anderweitig nicht genannt
9811 - Auszubildende mit (noch) nicht feststehendem Ausbildungs-beruf
9821 - Praktikanten/Praktikantinnen, Volontäre/ Volontärinnen mit (noch) nicht feststehendem Beruf
9911 - Facharbeiter/innen, ohne nähere Tätigkeitsangabe
9921 - Heimarbeiter/innen, ohne nähere Tätigkeitsangabe
9931 - Vorarbeiter/innen, Gruppenleiter/innen, ohne nähere Tätigkeitsangabe
9971 - Sonstige Arbeitskräfte, ohne nähere Tätigkeitsangabe
Statistisches Bundesamt (1996): Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 4.1.2., Beruf, Ausbildung und Arbeitsbedingung der Erwerbstätigen 1995 (Ergebnisse des Mikrozensus). Stuttgart: Metzler-Poeschel. pp. 317-323.
Hartmann/Schütz (2002): Die Klassifikation der Berufe und der Wirtschaftzweige im Sozio-oekonomischen Panel – Neuvercodung der Daten 1984 – 2001. Infratest Sozialforschung, München.

pgklas10 – Current Occupational Classification (KldB2010)

1104	[1104] Officer	7
1203	[1203] Senior Non-Commissioned Officers and Higher	4
1302	[1302] Junior Non-Commissioned Officers	1
1402	[1402] Armed Forces Personnel in Other Ranks	32
11101	[11101] Occupations in Farming (without Specialisation)-Unskilled/Semiskilled Tasks	8
11102	[11102] Occupations in Farming (without Specialisation)-Skilled Tasks	57
11103	[11103] Occupations in Farming (without Specialisation)-Complex Tasks	1
11104	[11104] Occupations in Farming (without Specialisation)-Highly Complex Tasks	1
11113	[11113] Technical Occup. in Farming-Complex Tasks	0
11114	[11114] Technical Occup. in Farming-Highly Complex Tasks	0
11123	[11123] Agricultural Experts-Complex Tasks	0
11124	[11124] Agricultural Experts-High Complex Tasks	0

11132	[11132] Technical Laboratory Occup. in Agriculture-Skilled Tasks	1
11133	[11133] Technical Laboratory Occup. in Agriculture-Complex Tasks	0
11182	[11182] Occupations in Farming (with Specialisation, Not Elsewhere Classified)-Skilled Tasks	0
...	(1263 rows omitted)	12630
94623	[94623] Prop Designers-Complex Tasks	0
94693	[94693] Supervisors in Stage, Costume and Prop Design	0
94704	[94704] Occupations in Museums (without Specialisation)-Highly Complex Tasks	1
94712	[94712] Technical Occup. in Museums and Exhibitions-Skilled Tasks	1
94713	[94713] Technical Occup. in Museums and Exhibitions-Complex Tasks	0
94714	[94714] Technical Occup. in Museums and Exhibitions-Highly Complex Tasks	0
94724	[94724] Art Experts-Highly Complex Tasks	0
94794	[94794] Managers in Museum	0
-1	[-1] No Answer	849
-2	[-2] Does not apply	10299
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	38398
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

This variable is designed to provide annual data on job classification for all employed persons according to the classification of the German Federal Statistical Office (StaBuA). Respondents answer the question on their current occupational title in their own words, and this response is entered into a blank in the questionnaire.
Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
In the SOEP-IS, pgklas10 was first implemented in wave 2016 and is coded [-5] in all previous waves.
German Classification of Occupations 2010. https://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen der-Berufe/KldB2010/Arbeitshilfen/EnglischeKldB2010/KldBEnglischl-Nav.html
Statistisches Bundesamt (1996): Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 4.1.2., Beruf, Ausbildung und Arbeitsbedingung der Erwerbstätigen 1995 (Ergebnisse des Mikrozensus). Stuttgart: Metzler-Poeschel. pp. 317-323.
Hartmann/Schütz (2002): Die Klassifikation der Berufe und der Wirtschaftzweige im Sozio-oekonomischen Panel – Neuvercodung der Daten 1984 – 2001. Infratest Sozialforschung, München.

pgautono – Autonomy in occupational activity

? In your position at work, do you supervise others? In other words, do people work under your direction? *(from: soep-is/soep-is-2019/Q320:pvor1[28340])*

0	[0] Apprentice	1305
1	[1] Low Autonomy	3299
2	[2] [2/5]	7830
3	[3] [3/5]	13038
4	[4] [4/5]	7645
5	[5] High Autonomy	1269
-1	[-1] No Answer	763

-2	[-2] Does Not Apply	27141
-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 gives the occupational autonomy for all employed persons. It offers an alternative to the ISCO-based scales on occupational status (PGISEI), class (PGEGP), or prestige (PGSIOPS). PGAUTONO is the simplest variable based on the scales of "occupational position" in terms of its construction, and strongly correlated with the Treiman Prestige Scale (PGSIOPS).
The basis for the "autonomy in occupational activity" scale is the classification of occupational position. Self-employed persons are categorized according to the size of the company (with the exception of farmers, who are all classified within the same category of autonomy, independent of farm size in hectares). Civil servants are differentiated according to the civil service laws defining each kind of activity and the amount of autonomy connected to it. Workers are differentiated according to their vocational training, and thus categorized hierarchically according to the different tasks they can be expected to carry out and the different amounts of responsibility associated with each task. Similarly, salaried employees are classified according to how differentiated their tasks are and how much responsibility is associated with each.
The value "1" is assigned mainly to manual workers with a low level of status and a low level of autonomy. Group 2 encompasses work in production, services demanding a minimal level of specialization, and farm work. Activities that require completion of the middle track of secondary education and entail a limited amount of responsibility are classified in Group 3. Group 4 includes activities carried out either with or without supervision that require a degree from a college of applied sciences or university, but are not very high in prestige. Managers and freelance academics are both placed in Group 5 (highest autonomy). Depending on the number of employees, self-employed are categorized in Group 3, Group 4, or Group 5.
Hoffmeyer-Zlotnik, Jürgen H.P., and Alfons J. Geis (2003) Berufs-klassifikation und Messung des beruflichen Status/ Prestige. In: ZUMA-Nachrichten 52, Jg. 27, Mai 2003. pp. 125-138.

pgisced – Highest degree/diploma attained, ISCED-1997

0	[0] (0) in school	750
1	[1] (1) inadequately	1066
2	[2] (2) general elemantary	8059
3	[3] (3) middle vocational	30338
4	[4] (4) vocational + Abi	4731
5	[5] (5) higher vocational	3729
6	[6] (6) higher education	12773
-1	[-1] No Answer	844
-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

To make the educational degrees and diplomas attained in different countries comparable, for all respondents an educational variable (PGISCED) is generated retroactively using the in-

ternational classification scheme ISCED-1997 (International Standard Classification of Education). It creates the highest degree/diploma attained, taking into account degrees and diplomas attained in both general schooling and in vocational and university education. Here the higher-level vocational and university override lower-level school diplomas. Persons who, for example, have no values for the variables on secondary school degrees/diplomas but state that they have a university degree are placed in the highest ISCED category. Please note that, due to a lack of more detailed information on tertiary degrees – in particular on doctoral degrees – we include all tertiary degrees in our ISCED category 6. Thus, the ISCED variable provided here is not comparable one-to-one with the ISCED levels as defined by the OECD, since we have included the original ISCED level 5A in our ISCED category 6. See below for more details.
Since the SOEP-IS does not include a youth questionnaire, the information usually coming from the youth questionnaire was not included in the generation of PGBBILA. Furthermore, since the year 2012, input information from PGBBILA is not being used as PGBBILA itself is not being generated.
OECD (1999): Classifying Educational Programmes: Manual for ISCED-97 Implementation in OECD Countries. Paris 1999.

pgcasmin – Highest degree/diploma according to CASMIN

0	[0] (0) In School	761
1	[1] (1a) Inadequately Completed	1051
2	[2] (1b) General Elementary School	4754
3	[3] (1c) Basic Vocational Qualification	15428
4	[4] (2b) Intermediate General Qualification	2627
5	[5] (2a) Intermediate Vocational	16227
6	[6] (2c_gen) General Maturity Certificate	2607
7	[7] (2c_voc) Vocational Maturity Certificate	5110
8	[8] (3a) Lower Tertiary Education	3537
9	[9] (3b) Higher Tertiary Education	9236
-1	[-1] No Answer	952
-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

As an alternative to PGISCED, a second educational variable is generated (PGCASMIN) that also enables comparison with international educational degrees/diplomas. Based on the modified CASMIN classification scheme (Comparative Analysis of Social Mobility in Industrial Nations), this variable has been computed retroactively from 1984 on for all respondents. Taken into account are both secondary-level and university/college-level degrees and diplomas. As with PGISCED, the higher-level occupational degrees override the lower-level secondary school degrees.
The original version is described in: König, W./Lüttinger, P./Müller, W. (1988): A Comparative Analysis of the Development and Structure of Educational Systems. Methodological Foundations and the Construction of a Comparative Educational Scale. CASMIN Working Paper No. 12. Mannheim: Universität Mannheim. For the modified version see: Brauns, H./Steinmann, (1999): Educational Reform in France, West-Germany and the United Kingdom: Updating the CASMIN Educational Classification. In: ZUMA Nachrichten, Jg. 23, H. 44, pp. 7-44.

pgstib – Occupational Position

? Are you currently enrolled in an educational or training program? In other words: are you in school or higher education, working on a doctors degree, completing vocational training, or taking part in further training? *(from: soep-is/soep-is-2019/Q291:paus1[28314])*

? What kind of general education / secondary school is it? *(from: soep-is/soep-is-2019/Q293:paus2[28316])*

? What is your current occupation? *(from: soep-is/soep-is-2019/Q313:pber[28333])*

? Which job position do you currently hold? *(from: soep-is/soep-is-2019/Q319:pstell[28339])*

? Which job position do you currently hold? // [Other position, namely:] *(from: soep-is/soep-is-2019/Q319:pstellso[28339])*

? In your position at work, do you supervise others? In other words, do people work under your direction? *(from: soep-is/soep-is-2019/Q320:pvor1[28340])*

? What is your current occupational status as a blue-collar worker? *(from: soep-is/soep-is-2019/Q392:parb[28410])*

? What is your current occupational status as a white-collar worker? *(from: soep-is/soep-is-2019/Q393:pang[28411])*

? What is your current occupational status as a civil servant? *(from: soep-is/soep-is-2019/Q394:pamt[28412])*

? What is your current occupational status as an apprentice / trainee or intern? *(from: soep-is/soep-is-2019/Q395:pazubi[28413])*

? In what occupational position are you currently self-employed? *(from: soep-is/soep-is-2019/Q428:psst[28444])*

? How many employees do you have? *(from: soep-is/soep-is-2019/Q429:psstanz[28445])*

0	[0] Do Not Know	0
10	[10] Not Employed	4320
11	[11] In Education	1941
12	[12] Unemployed, Not Employer	3165
13	[13] Pensioner	17603
15	[15] Military, Community Service	112
110	[110] Apprentice	43
120	[120] Apprentice,Trainee Industry Technology	675
130	[130] Apprentice, Trainee Trade And Commerce	438
140	[140] Trainee, Intern	149
150	[150] Aspirant	0
210	[210] Untrained Worker	956
220	[220] Semi-Trained Worker	2343
230	[230] Trained Worker	2465
240	[240] Foreman, Team Leader	393
...	(21 rows omitted)	6178
522	[522] Trained Employee With Simple Tasks	2957
530	[530] Qualified Professional	9846
540	[540] H. Qualified Professional	5263
550	[550] Managerial	494
610	[610] Low-Level Civil Service	76
620	[620] Middle-Level Civil Service	592
630	[630] High-Level Civil Service	969
640	[640] Executive Civil Service	555
999	[999] Employed Without StiB Info	0

-1	[-1] No Answer	756
-2	[-2] Does Not Apply	1
-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 variable represents a compilation of all relevant information on current occupational position. It is generated by combining information on "occupational group", "unemployed (yes/no)", "military/community service", "in education (yes/no)", and "pensioner". A hierarchical scheme is used to determine which data is given precedence when a variety of divergent information exists (increasing dominance):
10 – not employed
13 – pensioner
11 – currently in education
15 – military / community service
12 – registered unemployed
110-150 - apprentice
410-440 – self-employed
210-250 – manual laborer
510-550 - employee
610-640 – civil service
In PGSTIB, non-working persons are only assigned to the category (13) "pensioner" if they are recipients of retirement pension or if they are recipients of widow's pension AND are older than 60 years.

pgmonth – Month of interview

1	[1] January	4553
2	[2] February	5201
3	[3] March	3215
4	[4] April	1275
5	[5] May	1003
6	[6] June	726
7	[7] July	1266
8	[8] August	689
9	[9] September	11218
10	[10] October	19072
11	[11] November	10075
12	[12] December	3996
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	1
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

Month of interview is generated using the answers to the individual questionnaire. Missing answers are filled in using data from the HBRUTTO files.

pgmode – Interview method

100	[100] With Interviewer Assistance	30
110	[110] Oral Interview	2511
120	[120] Written Ques. Interviewer	2038
130	[130] Mix Between With/Without Interviewer	0
131	[131] Written Ques. No Interviewer	306
132	[132] Oral And Written	202
133	[133] Proxy	2
134	[134] With Interpreter	0
135	[135] Exc Interpreter	0
140	[140] CAPI - Wave O Onwards	57001
200	[200] Telephone Assistance	0
210	[210] Written, By Mail	181
220	[220] phone interview	0
300	[300] CAWI	18
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	1
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

The interview method is generated via the answers to the questions in the individual questionnaire. Missing answers are filled in from the PBRUTTO files.

pglabgro – Current gross labor income in euros (generated)

? How much did you earn from your work last month? If you had extra income in the last month, e.g.: vacation pay or subsequent payments, please do not accout that. But please include overtime payments. If you are self-employed, please estimate your monthly income before and after tax. If possible, please state both: your gross income, which means income before tax and social security deductions and your net income, which means income after tax, social security, and unemployment and health insurance deductions. // I earned gross: ... euros *(from: soep-is/soep-is-2019/Q434:pbrut[28450])*

0	261
1	2
8	1
14	1
15	1
20	2
22	1
25	2
30	2
35	1
40	9
45	2
48	1
50	19
51	1

...	(3142 rows omitted)	32479
22000		5
23000		1
24000		2
25000		2
30000		1
34000		1
40000		2
45000		1
50000		1
-1	[-1] No Answer	2133
-2	[-2] Does Not Apply	27314
-3	[-3] Answer Improbable	42
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

The variable PGLABGRO represents the current gross labor income of all SOEP respondents who are employed in each respective wave. The variable contains both generated and imputed values for Sample E (until 2011) & Sample I (until 2010); since then, the variable contains only generated values for all samples. The imputed values are available in the variables PGI1-PGI5LABGRO. Income details are consistently provided in euros for all waves. Item nonresponse is imputed in a twostage procedure: first, with the "Row-and-Column" method of Little und Su (1989) using individual longitudinal data as well as cross-sectional trend data (cf. Joachim R. Frick and Markus M. Grabka (2005): Item-Non-Response on Income Questions in Panel surveys: Incidence, Imputation and the Impact on the Income Distribution. Allgemeines Statistisches Archiv (ASTA) 89, 49-61). Alternatively, if no individual longitudinal information is available, we base the imputation on a regression using different Mincer covariates, also taking into account current net labor income. If both types of income information are lacking, first we impute current net labor income and then current gross labor income. Imputed values are flagged (PGIMPGRO).

pgi1labgro – 1. Imput. Akt. Bruttoerwerbseink.(gen) in Euro [1/5]

-1	[-1] No Answer	0
-2	[-2] Does Not Apply	21213
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	14520
-6	[-6] Questionnaire Version With Modified Filtering	0

Multiple imputation procedures provide a way to deal with missing values on the variable current net labor income in Euros by using information about determinants of the household income and replacing item-nonresponse with multiply imputed data. Five imputations are available within the $PGEN datasets: the variables pgi1labnet-pgi5labnet. The imputations were calculated using the method of chained equations predictive mean matching in STATA. The procedures were written by Patrick Royston (see Royston 2004, 2005a, 2005b, 2007, 2009) and Ian White (see White, Daniel and Royston 2010; White, Royston and Wood 2011). Predicted mean matching means that for each missing observation on income, the particular non-missing observation is found whose prediction on observed data is closest.

This closest observation is used to impute the missing value. The most important variable for modelling the current net labor income is the gross labor income of the previous year. A complete list of the variables used for modelling is available upon request. The missing observations were assumed to be missing at random. We set the number of imputations m=5 and get 5 multiple imputed values for pglabnet. The number of iterations carried out in each prediction model was specified to be 2000. Sample E&I and the supplementary sample S1 were imputed separately.
Analysing multiply imputed data: For analysing multiple imputed data, one does not necessarily need special methods; however, such tools exist and simplify the use of multiply imputed data. Below is given a short overview of some useful tools for various statistical packages. These tools estimate the parameters of a regression model by combining the estimates across the several replicates of imputation. Point estimates from multiple imputations are then the arithmetic mean of the several point estimates obtained from analysis on each imputed data. Standard errors are obtained by combining the average of the squared standard errors of the several (m) estimates with the within-and between-imputation variance.

- STATA provides a built-in functionality called mi.
- Within SAS, PROC MIANALYZE combines the results of analyses on the data sets.
- IVEware is a set of routines that can be launched from SAS or run independently using data from many sources. You can use the IVEware module regress to perform multiple imputation analysis.

Royston, P. 2004. Multiple imputation of missing values. Stata Journal 4: 227–241. Royston, P. 2005a. Multiple imputation of missing values: Update. Stata Journal 5: 188–201. Royston, P. 2005b. Multiple imputation of missing values: Update of ice. Stata Journal 5: 527–536. Royston, P. 2007. Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring. Stata Journal 7: 445–464. Royston, P. 2009. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. Stata Journal 9: 466–477.

pgi2labgro – 2. Imput. Akt. Bruttoerwerbseink.(gen) in Euro [2/5]

-1	[-1] No Answer	0
-2	[-2] Does Not Apply	21213
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	14520
-6	[-6] Questionnaire Version With Modified Filtering	0

Multiple imputation procedures provide a way to deal with missing values on the variable current net labor income in Euros by using information about determinants of the household income and replacing item-nonresponse with multiply imputed data. Five imputations are available within the $PGEN datasets: the variables pgi1labnet-pgi5labnet. The imputations were calculated using the method of chained equations predictive mean matching in STATA. The procedures were written by Patrick Royston (see Royston 2004, 2005a, 2005b, 2007, 2009) and Ian White (see White, Daniel and Royston 2010; White, Royston and Wood 2011). Predicted mean matching means that for each missing observation on income, the particular non-missing observation is found whose prediction on observed data is closest. This closest observation is used to impute the missing value. The most important variable for modelling the current net labor income is the gross labor income of the previous year. A complete list of the variables used for modelling is available upon request. The missing observations were assumed to be missing at random. We set the number of imputations m=5

and get 5 multiple imputed values for pglabnet. The number of iterations carried out in each prediction model was specified to be 2000. Sample E&I and the supplementary sample S1 were imputed separately.
Analysing multiply imputed data: For analysing multiple imputed data, one does not necessarily need special methods; however, such tools exist and simplify the use of multiply imputed data. Below is given a short overview of some useful tools for various statistical packages. These tools estimate the parameters of a regression model by combining the estimates across the several replicates of imputation. Point estimates from multiple imputations are then the arithmetic mean of the several point estimates obtained from analysis on each imputed data. Standard errors are obtained by combining the average of the squared standard errors of the several (m) estimates with the within-and between-imputation variance.

- STATA provides a built-in functionality called mi.
- Within SAS, PROC MIANALYZE combines the results of analyses on the data sets.
- IVEware is a set of routines that can be launched from SAS or run independently using data from many sources. You can use the IVEware module regress to perform multiple imputation analysis.

Royston, P. 2004. Multiple imputation of missing values. Stata Journal 4: 227–241. Royston, P. 2005a. Multiple imputation of missing values: Update. Stata Journal 5: 188–201. Royston, P. 2005b. Multiple imputation of missing values: Update of ice. Stata Journal 5: 527–536. Royston, P. 2007. Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring. Stata Journal 7: 445–464. Royston, P. 2009. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. Stata Journal 9: 466–477.

pgi3labgro – 3. Imput. Akt. Bruttoerwerbseink.(gen) in Euro [3/5]

-1	[-1] No Answer	0
-2	[-2] Does Not Apply	21213
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	14520
-6	[-6] Questionnaire Version With Modified Filtering	0

Multiple imputation procedures provide a way to deal with missing values on the variable current net labor income in Euros by using information about determinants of the household income and replacing item-nonresponse with multiply imputed data. Five imputations are available within the $PGEN datasets: the variables pgi1labnet-pgi5labnet. The imputations were calculated using the method of chained equations predictive mean matching in STATA. The procedures were written by Patrick Royston (see Royston 2004, 2005a, 2005b, 2007, 2009) and Ian White (see White, Daniel and Royston 2010; White, Royston and Wood 2011). Predicted mean matching means that for each missing observation on income, the particular non-missing observation is found whose prediction on observed data is closest. This closest observation is used to impute the missing value. The most important variable for modelling the current net labor income is the gross labor income of the previous year. A complete list of the variables used for modelling is available upon request. The missing observations were assumed to be missing at random. We set the number of imputations m=5 and get 5 multiple imputed values for pglabnet. The number of iterations carried out in each prediction model was specified to be 2000. Sample E&I and the supplementary sample S1 were imputed separately.
Analysing multiply imputed data: For analysing multiple imputed data, one does not ne-

cessarily need special methods; however, such tools exist and simplify the use of multiply imputed data. Below is given a short overview of some useful tools for various statistical packages. These tools estimate the parameters of a regression model by combining the estimates across the several replicates of imputation. Point estimates from multiple imputations are then the arithmetic mean of the several point estimates obtained from analysis on each imputed data. Standard errors are obtained by combining the average of the squared standard errors of the several (m) estimates with the within-and between-imputation variance.

- STATA provides a built-in functionality called mi.
- Within SAS, PROC MIANALYZE combines the results of analyses on the data sets.
- IVEware is a set of routines that can be launched from SAS or run independently using data from many sources. You can use the IVEware module regress to perform multiple imputation analysis.

Royston, P. 2004. Multiple imputation of missing values. Stata Journal 4: 227–241. Royston, P. 2005a. Multiple imputation of missing values: Update. Stata Journal 5: 188–201. Royston, P. 2005b. Multiple imputation of missing values: Update of ice. Stata Journal 5: 527–536. Royston, P. 2007. Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring. Stata Journal 7: 445–464. Royston, P. 2009. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. Stata Journal 9: 466–477.

pgi4labgro – 4. Imput. Akt. Bruttoerwerbseink.(gen) in Euro [4/5]

-1	[-1] No Answer	0
-2	[-2] Does Not Apply	21213
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	14520
-6	[-6] Questionnaire Version With Modified Filtering	0

Multiple imputation procedures provide a way to deal with missing values on the variable current net labor income in Euros by using information about determinants of the household income and replacing item-nonresponse with multiply imputed data. Five imputations are available within the $PGEN datasets: the variables pgi1labnet-pgi5labnet. The imputations were calculated using the method of chained equations predictive mean matching in STATA. The procedures were written by Patrick Royston (see Royston 2004, 2005a, 2005b, 2007, 2009) and Ian White (see White, Daniel and Royston 2010; White, Royston and Wood 2011). Predicted mean matching means that for each missing observation on income, the particular non-missing observation is found whose prediction on observed data is closest. This closest observation is used to impute the missing value. The most important variable for modelling the current net labor income is the gross labor income of the previous year. A complete list of the variables used for modelling is available upon request. The missing observations were assumed to be missing at random. We set the number of imputations m=5 and get 5 multiple imputed values for pglabnet. The number of iterations carried out in each prediction model was specified to be 2000. Sample E&I and the supplementary sample S1 were imputed separately.

Analysing multiply imputed data: For analysing multiple imputed data, one does not necessarily need special methods; however, such tools exist and simplify the use of multiply imputed data. Below is given a short overview of some useful tools for various statistical packages. These tools estimate the parameters of a regression model by combining the estimates across the several replicates of imputation. Point estimates from multiple imputations

are then the arithmetic mean of the several point estimates obtained from analysis on each imputed data. Standard errors are obtained by combining the average of the squared standard errors of the several (m) estimates with the within-and between-imputation variance.

- STATA provides a built-in functionality called mi.
- Within SAS, PROC MIANALYZE combines the results of analyses on the data sets.
- IVEware is a set of routines that can be launched from SAS or run independently using data from many sources. You can use the IVEware module regress to perform multiple imputation analysis.

Royston, P. 2004. Multiple imputation of missing values. Stata Journal 4: 227–241. Royston, P. 2005a. Multiple imputation of missing values: Update. Stata Journal 5: 188–201. Royston, P. 2005b. Multiple imputation of missing values: Update of ice. Stata Journal 5: 527–536. Royston, P. 2007. Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring. Stata Journal 7: 445–464. Royston, P. 2009. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. Stata Journal 9: 466–477.

pgi5labgro – 5. Imput. Akt. Bruttoerwerbseink.(gen) in Euro [5/5]

-1	[-1] No Answer	0
-2	[-2] Does Not Apply	21213
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	14520
-6	[-6] Questionnaire Version With Modified Filtering	0

Multiple imputation procedures provide a way to deal with missing values on the variable current net labor income in Euros by using information about determinants of the household income and replacing item-nonresponse with multiply imputed data. Five imputations are available within the $PGEN datasets: the variables pgi1labnet-pgi5labnet. The imputations were calculated using the method of chained equations predictive mean matching in STATA. The procedures were written by Patrick Royston (see Royston 2004, 2005a, 2005b, 2007, 2009) and Ian White (see White, Daniel and Royston 2010; White, Royston and Wood 2011). Predicted mean matching means that for each missing observation on income, the particular non-missing observation is found whose prediction on observed data is closest. This closest observation is used to impute the missing value. The most important variable for modelling the current net labor income is the gross labor income of the previous year. A complete list of the variables used for modelling is available upon request. The missing observations were assumed to be missing at random. We set the number of imputations m=5 and get 5 multiple imputed values for pglabnet. The number of iterations carried out in each prediction model was specified to be 2000. Sample E&I and the supplementary sample S1 were imputed separately.

Analysing multiply imputed data: For analysing multiple imputed data, one does not necessarily need special methods; however, such tools exist and simplify the use of multiply imputed data. Below is given a short overview of some useful tools for various statistical packages. These tools estimate the parameters of a regression model by combining the estimates across the several replicates of imputation. Point estimates from multiple imputations are then the arithmetic mean of the several point estimates obtained from analysis on each imputed data. Standard errors are obtained by combining the average of the squared standard errors of the several (m) estimates with the within-and between-imputation variance.

- STATA provides a built-in functionality called mi.

· Within SAS, PROC MIANALYZE combines the results of analyses on the data sets.
· IVEware is a set of routines that can be launched from SAS or run independently using data from many sources. You can use the IVEware module regress to perform multiple imputation analysis.
Royston, P. 2004. Multiple imputation of missing values. Stata Journal 4: 227–241. Royston, P. 2005a. Multiple imputation of missing values: Update. Stata Journal 5: 188–201. Royston, P. 2005b. Multiple imputation of missing values: Update of ice. Stata Journal 5: 527–536. Royston, P. 2007. Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring. Stata Journal 7: 445–464. Royston, P. 2009. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. Stata Journal 9: 466–477.

pgimpgro – Imputation flag for LABGROxx

0	[0] not imputed	31097
1	[1] Imputed	3879
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	27314
-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 variable PGIMPGRO designates imputations of item-nonresponse in the variable PGLABGRO (current gross labor income).

pglabnet – Current net labor income (generated) in euros

? How much did you earn from your work last month? If you had extra income in the last month, e.g.: vacation pay or subsequent payments, please do not accout that. But please include overtime payments. If you are self-employed, please estimate your monthly income before and after tax. If possible, please state both: your gross income, which means income before tax and social security deductions and your net income, which means income after tax, social security, and unemployment and health insurance deductions. // I earned net: ... euros *(from: soep-is/soep-is-2019/Q434:pnett[28450])*

-1	[-1] No Answer	1537
-2	[-2] Does Not Apply	27314
-3	[-3] Answer Improbable	32
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

The variable PGLABNET represents the current net labor income of all persons in SOEP working in the respective wave. The variable contains both generated and imputed values for Sample E (until 2011) & Sample I (until 2010); since then, the variable contains only generated values for all samples. The imputed values are available in the variables PGI1-PGI5LABNET. Income details are consistently provided in euros for all waves. The imputation of item nonresponse takes place in a two-stage procedure: first, with the "Row-and-Column" method of Little und Su (1989) using individual longitudinal data as well as cross-sectional trend data

(cf. Joachim R. Frick and Markus M. Grabka (2005): Item-Non-Response on Income Questions in Panel surveys: Incidence, Imputation and the Impact on the Income Distribution. Allgemeines Statistisches Archiv (ASTA) 89, 49-61). Alternatively, if no individual longitudinal information is available, we base the imputation on a regression using different Mincer covariates, also taking into account current gross labor income. If both types of income information are lacking, first we impute current gross labor income and then current net labor income. Imputed values are flagged (PGIMPNET).

pgi1labnet – 1. Imput. Akt. Nettoerwerbseink.(gen) in Euro [1/5]

-1	[-1] No Answer	0
-2	[-2] Does Not Apply	21213
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	14520
-6	[-6] Questionnaire Version With Modified Filtering	0

Multiple imputation procedures provide a way to deal with missing values on the variable current net labor income in Euros by using information about determinants of the household income and replacing item-nonresponse with multiply imputed data. Five imputations are available within the $PGEN datasets: the variables pgi1labnet-pgi5labnet. The imputations were calculated using the method of chained equations predictive mean matching in STATA. The procedures were written by Patrick Royston (see Royston 2004, 2005a, 2005b, 2007, 2009) and Ian White (see White, Daniel and Royston 2010; White, Royston and Wood 2011). Predicted mean matching means that for each missing observation on income, the particular non-missing observation is found whose prediction on observed data is closest. This closest observation is used to impute the missing value. The most important variable for modelling the current net labor income is the gross labor income of the previous year. A complete list of the variables used for modelling is available upon request. The missing observations were assumed to be missing at random. We set the number of imputations m=5 and get 5 multiple imputed values for pglabnet. The number of iterations carried out in each prediction model was specified to be 2000. Sample E&I and the supplementary sample S1 were imputed separately.
Analysing multiply imputed data: For analysing multiple imputed data, one does not necessarily need special methods; however, such tools exist and simplify the use of multiply imputed data. Below is given a short overview of some useful tools for various statistical packages. These tools estimate the parameters of a regression model by combining the estimates across the several replicates of imputation. Point estimates from multiple imputations are then the arithmetic mean of the several point estimates obtained from analysis on each imputed data. Standard errors are obtained by combining the average of the squared standard errors of the several (m) estimates with the within-and between-imputation variance.

- STATA provides a built-in functionality called mi.
- Within SAS, PROC MIANALYZE combines the results of analyses on the data sets.
- IVEware is a set of routines that can be launched from SAS or run independently using data from many sources. You can use the IVEware module regress to perform multiple imputation analysis.

Royston, P. 2004. Multiple imputation of missing values. Stata Journal 4: 227–241. Royston, P. 2005a. Multiple imputation of missing values: Update. Stata Journal 5: 188–201. Royston, P. 2005b. Multiple imputation of missing values: Update of ice. Stata Journal 5: 527–536. Royston, P. 2007. Multiple imputation of missing values: Further update of ice, with an emphasis

on interval censoring. Stata Journal 7: 445–464. Royston, P. 2009. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. Stata Journal 9: 466–477.

pgi2labnet – 2. Imput. Akt. Nettoerwerbseink.(gen) in Euro [2/5]

-1	[-1] No Answer	0
-2	[-2] Does Not Apply	21213
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	14520
-6	[-6] Questionnaire Version With Modified Filtering	0

Multiple imputation procedures provide a way to deal with missing values on the variable current net labor income in Euros by using information about determinants of the household income and replacing item-nonresponse with multiply imputed data. Five imputations are available within the $PGEN datasets: the variables pgi1labnet-pgi5labnet. The imputations were calculated using the method of chained equations predictive mean matching in STATA. The procedures were written by Patrick Royston (see Royston 2004, 2005a, 2005b, 2007, 2009) and Ian White (see White, Daniel and Royston 2010; White, Royston and Wood 2011). Predicted mean matching means that for each missing observation on income, the particular non-missing observation is found whose prediction on observed data is closest. This closest observation is used to impute the missing value. The most important variable for modelling the current net labor income is the gross labor income of the previous year. A complete list of the variables used for modelling is available upon request. The missing observations were assumed to be missing at random. We set the number of imputations m=5 and get 5 multiple imputed values for pglabnet. The number of iterations carried out in each prediction model was specified to be 2000. Sample E&I and the supplementary sample S1 were imputed separately.
Analysing multiply imputed data: For analysing multiple imputed data, one does not necessarily need special methods; however, such tools exist and simplify the use of multiply imputed data. Below is given a short overview of some useful tools for various statistical packages. These tools estimate the parameters of a regression model by combining the estimates across the several replicates of imputation. Point estimates from multiple imputations are then the arithmetic mean of the several point estimates obtained from analysis on each imputed data. Standard errors are obtained by combining the average of the squared standard errors of the several (m) estimates with the within-and between-imputation variance.

- STATA provides a built-in functionality called mi.
- Within SAS, PROC MIANALYZE combines the results of analyses on the data sets.
- IVEware is a set of routines that can be launched from SAS or run independently using data from many sources. You can use the IVEware module regress to perform multiple imputation analysis.

Royston, P. 2004. Multiple imputation of missing values. Stata Journal 4: 227–241. Royston, P. 2005a. Multiple imputation of missing values: Update. Stata Journal 5: 188–201. Royston, P. 2005b. Multiple imputation of missing values: Update of ice. Stata Journal 5: 527–536. Royston, P. 2007. Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring. Stata Journal 7: 445–464. Royston, P. 2009. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. Stata Journal 9: 466–477.

pgi3labnet – 3. Imput. Akt. Nettoerwerbseink.(gen) in Euro [3/5]

-1	[-1] No Answer	0
-2	[-2] Does Not Apply	21213
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	14520
-6	[-6] Questionnaire Version With Modified Filtering	0

Multiple imputation procedures provide a way to deal with missing values on the variable current net labor income in Euros by using information about determinants of the household income and replacing item-nonresponse with multiply imputed data. Five imputations are available within the $PGEN datasets: the variables pgi1labnet-pgi5labnet. The imputations were calculated using the method of chained equations predictive mean matching in STATA. The procedures were written by Patrick Royston (see Royston 2004, 2005a, 2005b, 2007, 2009) and Ian White (see White, Daniel and Royston 2010; White, Royston and Wood 2011). Predicted mean matching means that for each missing observation on income, the particular non-missing observation is found whose prediction on observed data is closest. This closest observation is used to impute the missing value. The most important variable for modelling the current net labor income is the gross labor income of the previous year. A complete list of the variables used for modelling is available upon request. The missing observations were assumed to be missing at random. We set the number of imputations m=5 and get 5 multiple imputed values for pglabnet. The number of iterations carried out in each prediction model was specified to be 2000. Sample E&I and the supplementary sample S1 were imputed separately.
Analysing multiply imputed data: For analysing multiple imputed data, one does not necessarily need special methods; however, such tools exist and simplify the use of multiply imputed data. Below is given a short overview of some useful tools for various statistical packages. These tools estimate the parameters of a regression model by combining the estimates across the several replicates of imputation. Point estimates from multiple imputations are then the arithmetic mean of the several point estimates obtained from analysis on each imputed data. Standard errors are obtained by combining the average of the squared standard errors of the several (m) estimates with the within-and between-imputation variance.

- STATA provides a built-in functionality called mi.
- Within SAS, PROC MIANALYZE combines the results of analyses on the data sets.
- IVEware is a set of routines that can be launched from SAS or run independently using data from many sources. You can use the IVEware module regress to perform multiple imputation analysis.

Royston, P. 2004. Multiple imputation of missing values. Stata Journal 4: 227–241. Royston, P. 2005a. Multiple imputation of missing values: Update. Stata Journal 5: 188–201. Royston, P. 2005b. Multiple imputation of missing values: Update of ice. Stata Journal 5: 527–536. Royston, P. 2007. Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring. Stata Journal 7: 445–464. Royston, P. 2009. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. Stata Journal 9: 466–477.

pgi4labnet – 4. Imput. Akt. Nettoerwerbseink.(gen) in Euro [4/5]

-1	[-1] No Answer	0
-2	[-2] Does Not Apply	21213

-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	14520
-6	[-6] Questionnaire Version With Modified Filtering	0

Multiple imputation procedures provide a way to deal with missing values on the variable current net labor income in Euros by using information about determinants of the household income and replacing item-nonresponse with multiply imputed data. Five imputations are available within the $PGEN datasets: the variables pgi1labnet-pgi5labnet. The imputations were calculated using the method of chained equations predictive mean matching in STATA. The procedures were written by Patrick Royston (see Royston 2004, 2005a, 2005b, 2007, 2009) and Ian White (see White, Daniel and Royston 2010; White, Royston and Wood 2011). Predicted mean matching means that for each missing observation on income, the particular non-missing observation is found whose prediction on observed data is closest. This closest observation is used to impute the missing value. The most important variable for modelling the current net labor income is the gross labor income of the previous year. A complete list of the variables used for modelling is available upon request. The missing observations were assumed to be missing at random. We set the number of imputations m=5 and get 5 multiple imputed values for pglabnet. The number of iterations carried out in each prediction model was specified to be 2000. Sample E&I and the supplementary sample S1 were imputed separately.
Analysing multiply imputed data: For analysing multiple imputed data, one does not necessarily need special methods; however, such tools exist and simplify the use of multiply imputed data. Below is given a short overview of some useful tools for various statistical packages. These tools estimate the parameters of a regression model by combining the estimates across the several replicates of imputation. Point estimates from multiple imputations are then the arithmetic mean of the several point estimates obtained from analysis on each imputed data. Standard errors are obtained by combining the average of the squared standard errors of the several (m) estimates with the within-and between-imputation variance.

- STATA provides a built-in functionality called mi.
- Within SAS, PROC MIANALYZE combines the results of analyses on the data sets.
- IVEware is a set of routines that can be launched from SAS or run independently using data from many sources. You can use the IVEware module regress to perform multiple imputation analysis.

Royston, P. 2004. Multiple imputation of missing values. Stata Journal 4: 227–241. Royston, P. 2005a. Multiple imputation of missing values: Update. Stata Journal 5: 188–201. Royston, P. 2005b. Multiple imputation of missing values: Update of ice. Stata Journal 5: 527–536. Royston, P. 2007. Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring. Stata Journal 7: 445–464. Royston, P. 2009. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. Stata Journal 9: 466–477.

pgi5labnet – 5. Imput. Akt. Nettoerwerbseink.(gen) in Euro [5/5]

-1	[-1] No Answer	0
-2	[-2] Does Not Apply	21213
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	14520
-6	[-6] Questionnaire Version With Modified Filtering	0

Multiple imputation procedures provide a way to deal with missing values on the variable current net labor income in Euros by using information about determinants of the household income and replacing item-nonresponse with multiply imputed data. Five imputations are available within the $PGEN datasets: the variables pgi1labnet-pgi5labnet. The imputations were calculated using the method of chained equations predictive mean matching in STATA. The procedures were written by Patrick Royston (see Royston 2004, 2005a, 2005b, 2007, 2009) and Ian White (see White, Daniel and Royston 2010; White, Royston and Wood 2011). Predicted mean matching means that for each missing observation on income, the particular non-missing observation is found whose prediction on observed data is closest. This closest observation is used to impute the missing value. The most important variable for modelling the current net labor income is the gross labor income of the previous year. A complete list of the variables used for modelling is available upon request. The missing observations were assumed to be missing at random. We set the number of imputations m=5 and get 5 multiple imputed values for pglabnet. The number of iterations carried out in each prediction model was specified to be 2000. Sample E&I and the supplementary sample S1 were imputed separately.
Analysing multiply imputed data: For analysing multiple imputed data, one does not necessarily need special methods; however, such tools exist and simplify the use of multiply imputed data. Below is given a short overview of some useful tools for various statistical packages. These tools estimate the parameters of a regression model by combining the estimates across the several replicates of imputation. Point estimates from multiple imputations are then the arithmetic mean of the several point estimates obtained from analysis on each imputed data. Standard errors are obtained by combining the average of the squared standard errors of the several (m) estimates with the within-and between-imputation variance.

- STATA provides a built-in functionality called mi.
- Within SAS, PROC MIANALYZE combines the results of analyses on the data sets.
- IVEware is a set of routines that can be launched from SAS or run independently using data from many sources. You can use the IVEware module regress to perform multiple imputation analysis.

Royston, P. 2004. Multiple imputation of missing values. Stata Journal 4: 227–241. Royston, P. 2005a. Multiple imputation of missing values: Update. Stata Journal 5: 188–201. Royston, P. 2005b. Multiple imputation of missing values: Update of ice. Stata Journal 5: 527–536. Royston, P. 2007. Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring. Stata Journal 7: 445–464. Royston, P. 2009. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. Stata Journal 9: 466–477.

pgimpnet – Imputation flag for LABNETxx

0	[0] not imputed	32093
1	[1] Imputed	2883
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	27314
-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 variable PGIMPNET designates imputations of item-nonresponse in the variable PGLABNET (current net labor income).

pgallbet – Core size category of the company

1	[1] Less Than 20	6858
2	[2] 20 Up To 200	8309
3	[3] 200 Up To 2000	6250
4	[4] 2000 And More	8789
5	[5] Self-Employed Without Coworkers	0
-1	[-1] No Answer	1308
-2	[-2] Does Not Apply	30770
-3	[-3] Answer Improbable	6
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	0
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable is designed to provide annual data on the core size category of the company for all employed persons.
Not all employed persons are asked the relevant input questions on an annual basis. Only those employed persons who changed jobs and first-time respondents were asked to provide up-to-date information.
Self-employed are not included in this variable. Information about the company size is included in the variable pgstib.

pgemplst – Employment status

? Are you currently employed? Which one of the following applies best to your status? *(from: soep-is/soep-is-2019/Q302:perw[28325])*

1	[1] Full-Time Employment	22105
2	[2] Regular Part-Time	8120
3	[3] Vocational Training	1237
4	[4] Marginal, Irregular Part-Time Employment	3356
5	[5] Not Employed	27363
6	[6] Sheltered workshop	108
-1	[-1] No Answer	1
-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

This variable is generated from the annual question on current employment status, which has a central filter function in the questionnaire to separate employed people from non-employed people for further questions. It is designed to provide consistent longitudinal data on employment status across all waves.
The category "not employed" comprises non-working individuals, those in military/community service, those on maternity leave, and employed persons in a phased retirement scheme (Altersteilzeit) whose current actual working hours are zero.
PGEMPLST supplements the variable PGLFS, which differentiates among persons who are not employed.

pgexpft – Working experience full-time employment

0		1675
0.1		18
0.2		28
0.3		49
0.4		20
0.5		196
0.6		21
0.7		20
0.8		36
0.9		36
1		144
1.1		20
1.2		22
1.3		22
1.4		26
...	(448 rows omitted)	12106
47.4		1
47.8		11
48		11
48.1		2
48.2		1
48.3		1
48.6		11
49.3		2
50		6
-1	[-1] No Answer	840
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	46965
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable reflects the total length of full-time employment in the respondent's career. The variable is created by combining monthly information on employment status from the calendar dataset ARTKALEN (which provides monthly information on activity status since an individual entered the SOEP) and annual information from the biographical dataset PBIOSPE (which provides information on activity status over the individual's life course). PGEXPFT uses calendar information up to December of the previous year and gives the length of time in years with months in decimal form.
If there is no monthly calendar data available in a given year of a respondent's career, the annual data from PBIOSPE is used for that year. In the most current wave the variable only uses up-to-date information from the newly answered Biography Questionnaires. If the year in which a spell started and ended is the same, and if there is no monthly data, a spell of 0.5 years is assumed.
Persons whose life course has been observed completely but with no spell of full-time employment are assigned the code (0). The code (-1) is assigned to all persons whose life course has not been observed completely. Persons with inconsistent information receive a (-3).
Because detailed information on the acitivity status of the respondents is not assessed in the

Questionnaire of the SOEP Innovation Sample, PGEXPFT is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGEXPFT is coded to "-5" (not contained in questionnaire).
Please also see PGEXPPT and PGEXPUE.

pgexppt – Working experience part-time employment

0		9031
0.1		70
0.2		117
0.3		112
0.4		76
0.5		249
0.6		67
0.7		67
0.8		91
0.9		28
1		436
1.1		57
1.2		62
1.3		97
1.4		37
...	(276 rows omitted)	3862
38		4
38.2		1
39.2		1
40		14
40.2		1
41.3		1
42		2
42.2		1
43		1
-1	[-1] No Answer	840
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	46965
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable reflects the total length of part-time employment in the respondent's career. The variable is created by combining monthly information on employment status from the calendar dataset ARTKALEN (which provides monthly information on activity status since an individual entered the SOEP) and annual information from the biographical dataset PBIOSPE (which provides information on activity status over the life course of an individual). PGEXPPT uses calendar information up to December of the previous year and gives the length of time in years with months in decimal form.
If there is no monthly calendar data available in a given year of a respondent's career, the annual data from PBIOSPE is used for that year. In the most current wave the variable only uses up-to-date information from the newly answered Biography Questionnaires. If the year in which a spell started and ended is the same, and if there is no monthly data, a spell of 0.5

years is assumed.
Persons whose life course has been observed completely but with no spell of full-time employment are assigned the code (0). The code (-1) is assigned to all persons whose life course has not been observed completely. Persons with inconsistent information receive a (-3).
Because detailed information on the activity status of the respondents is not assessed in the Questionnaire of the SOEP Innovation Sample, PGEXPPT is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGEXPPT is coded to "-5" (not contained in questionnaire).
Please also see PGEXPFT and PGEXPUE.

pgexpue – Unemployment experience

0		10635
0.1		149
0.2		123
0.3		173
0.4		83
0.5		531
0.6		72
0.7		91
0.8		128
0.9		66
1		409
1.1		103
1.2		39
1.3		59
1.4		102
...	(146 rows omitted)	1712
22.5		1
23		1
23.5		1
24		1
24.5		1
25.3		1
26.4		2
32		1
33		1
-1	[-1] No Answer	840
-2	[-2] Does Not Apply	0
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	46965
-6	[-6] Questionnaire Version With Modified Filtering	0

This variable reflects the total length of unemployment in the respondent's career. The variable is created by combining monthly information on employment status from the calendar dataset ARTKALEN (which provides monthly information on activity status since an individual entered the SOEP) and annual information from the biographical dataset PBIOSPE (which provides information on activity status over the life course of an individual). PGEXPUE uses calendar information up to December of the previous year and gives the length of

time in years with months in decimal form.
If there is no monthly calendar data available on a given year in a respondent's career, the annual data from PBIOSPE is used for that year. In the most current wave the variable only uses up-to-date information from the newly answered Biography Questionnaires. If the year in which a spell started and ended is the same, and if there is no monthly data, a spell of 0.5 years is assumed.
Persons whose life course has been observed completely but with no spell of full-time employment are assigned the code (0). The code (-1) is assigned to all persons whose life course has not been observed completely. Persons with inconsistent information receive a (-3).
Because detailed information on the activity status of the respondents is not assessed in the Questionnaire of the SOEP Innovation Sample, PGEXPUE is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGEXPUE is coded to "-5" (not contained in questionnaire).
Please also see PGEXPFT.

pgjobch – Occupational Change

1	[1] Not Employed	27141
2	[2] Employed No Change	24480
3	[3] Employed No Info If Change	3119
4	[4] Employed With Change	6166
5	[5] First Job	748
-1	[-1] No Answer	636
-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

This variable indicates a change of job since the previous interview for respondents with a follow-up interview, whereas for first-time respondents, the information refers to a change of job since the beginning of the previous year.
PGJOBCH is generated based on the central filter variable, which indicates whether a respondent has changed jobs since the beginning of the previous year. Information on the date of job change is then combined with interview month of the previous year's interview to identify whether a new job change has taken place since the previous interview.
Hence, PGJOBCH is a modified version of the variable PGERWTYP which indicates whether a respondent has changed jobs since the beginning of the previous year. Unlike PGERWTYP, the variable is calculated for all waves, and the codes are assigned independently of the respondent being a first-time or follow-up respondent.
In addition to PGERWTYP, the variable is also designed to identify respondents who have entered employment for the first time.
In addition to PGERWTYP, the variable is designed to provide consistent longitudinal information on job changes. The PGJOBCH variable is generated by correcting the original job change information in various ways:

1. We check whether the job changes stated by a respondent in two consecutives interviews refer to one and the same job change. The date of the job change and the interview month are used to correct double entries.

2. If the respondent indicates a job change with a date before the previous interview but did not state a job change in the previous interview, this is coded as a job change in the current interview.

3. If a respondent indicates no job change and was not employed at the time of the previous interview, this is coded as "no job change" despite the seeming implausibility, since there are possible explanations how this information could be plausible, e.g. if there were short-term employment spells between two interview dates.

4. Respondents can be "first-time employed" only once. If a respondent states being "first-time employed" for a second time, this is coded as "employed, with change".

pgfield – Field of tertiary education

1	[1] Applied Linguistics and Cultural Studies	0
2	[2] Protestant Theology	3
3	[3] Catholic Theology	0
4	[4] Philosophy	0
5	[5] History	2
6	[6] Library Science, Archival Studies, Journalism	12
7	[7] Literary Studies, Linguistics	4
8	[8] Classical Philology, Modern Greek	0
9	[9] German Philology	12
10	[10] English Studies	12
11	[11] Roman Studies	0
12	[12] Slavonic Studies	0
13	[13] Non-European Languages and Cultural Studies	0
14	[14] Cultural Studies	2
15	[15] Psychology	5
...	(36 rows omitted)	537
68	[68] Civil Engineering	19
69	[69] Surveying and Mapping	0
74	[74] Art, Aesthetics	0
75	[75] Fine Arts	0
76	[76] Design	0
77	[77] Performance, Film and Television, Theater	0
78	[78] Music, Musicology	7
83	[83] Outside the structure of the university system	0
98	[98] Not categorizable	57
-1	[-1] No Answer	70
-2	[-2] Does Not Apply	14582
-3	[-3] Answer Improbable	1
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	46965
-6	[-6] Questionnaire Version With Modified Filtering	0

The variable is designed to provide information on the field of education of tertiary degrees which adds details to the information recorded in the variable PGBBIL02. While the latter variable records if a person holds a degree PGFIELD contains more detailed information on the type of the degree. The data of the generated variable PGFIELD stem from two sources: 1. Person questionnaire: Each year since 1985 respondents are asked if they have left education

since the beginning of the year prior to the survey and which degrees they have obtained. This part of the questionnaire contains an open question on the type and the field of newly obtained tertiary degrees. This information is coded and used for the generation of the variables PGFIELD. 2. Biography questionnaire: Since 2001 similar information is collected from respondents who fill in the biography questionnaire (usually during the first two years of participation in the panel). In contrast to the information from the person questionnaire the questions do not refer to currently obtained degrees but to degrees obtained during the time before being part of the SOEP sample.
In the variable PGFIELD we combine these two types of information. Each year the variable contains the most recently collected information.
If you want to take into account that a person holds two degrees you have to combine the information from all available years. However, only a minority of the population holds more than one tertiary degree. In very few cases we encounter the problem that a respondent provides information on two different degrees in one survey year. This only happens in years when respondents fill in the person as well as the biography questionnaire. In these cases we prioritize the information from the person questionnaire as it refers to the current situation while the biography questionnaire contains retrospective information. Furthermore, there are cases who report an applied university degree and a university degree in the biography questionnaire. In these cases, the variable contains information on the university degree only.
The variable is coded according to the classification on fields of education ("Fächergruppen") provided by the Statistisches Bundesamt (2009). Until 2009 data from the person questionnaire were coded using an earlier version of this classification (1982). In the variable PGFIELD we recoded the original values. As the newer version is more precise this could be done with hardly any loss of information. Some categories are collapsed. Category 3 is coded as 2 (no distinction between catholic and protestant theology), 14 as 13, 17 as 16, 24 as 23, 25 as 26 and 48 as 49.
Because detailed information on the field of tertiary education is not assessed in the Questionnaire of the SOEP Innovation Sample, PGFIELD is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGFIELD is coded to "-5" (not contained in questionnaire).
Stat. Bundesamt (2009): Bildung und Kultur. Studierende an Hochschulen, Fachserie 11 Reihe 4.1, Wiesbaden: 446ff, Übersicht 1: „Fächergruppen, Studienbereiche und Studienfächer".

pgdegree – Type of tertiary degree

11	[11] Magister	12
12	[12] Diplom (University)	164
13	[13] Bachelor	6
14	[14] Master	2
15	[15] 1st State Examination	13
16	[16] Other state examination	14
21	[21] Diplom (at technical college, technical college for administration)	184
22	[22] Bachelor (at technical college, technical college for administration)	8
23	[23] Master (at technical college, technical college for administration)	1
31	[31] Teacher training,BA,MA at elementary, lower secondary schools/primary level	36
32	[32] Teacher training,BA,MA at 2ndary level 1/elementary schools/primary level	0
33	[33] Teacher training,BA,MA at intermediate scndry schools/scndry level I	5
34	[34] Teacher training, BA, MA at secondary level II and I	0
35	[35] Teacher training,BA,MA at academic 2ndry schools,2ndry levl 2,genrl school	15

36	[36] Teacher training, BA, MA at special needs schools	3
37	[37] Teacher training, BA, MA at vocational schools	5
38	[38] Teacher training, other	28
41	[41] Academic degree in the arts	1
42	[42] Doctorate	19
43	[43] Post-doctoral dissertation (Habilitation)	0
44	[44] Other Degree	8
98	[98] Not categorizable	98
-1	[-1] No Answer	120
-2	[-2] Does Not Apply	14582
-3	[-3] Answer Improbable	1
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	46965
-6	[-6] Questionnaire Version With Modified Filtering	0

The variable is designed to provide information on the type of tertiary degree (e.g., Diploma, Bachelor, Master) which adds details to the information recorded in the variable PGBBIL02. While the latter variable records if a person holds a degree PGDEGREE contains more detailed information on the type of the degree. The data of the generated variable PGDEGREE stem from two sources: 1. Person questionnaire: Each year since 1985 respondents are asked if they have left education since the beginning of the year prior to the survey and which degrees they have obtained. This part of the questionnaire contains an open question on the type and the field of newly obtained tertiary degrees. This information is coded and used for the generation of the variables PGDEGREE. 2. Biography questionnaire: Since 2001 similar information is collected from respondents who fill in the biography questionnaire (usually during the first two years of participation in the panel). In contrast to the information from the person questionnaire the questions do not refer to currently obtained degrees but to degrees obtained during the time before being part of the SOEP sample.
In the variable PGDEGREE we combine these two types of information. However, since the retrospective information was not collected before 2001 the variable covers until 2000 only persons for whom we have prospectively observed the end of study. This explains why the number of valid observations is rather small in these years.
Each year the variable contains the most recently collected information. If you want to take into account that a person holds two degrees you have to combine the information from all available years. However, only a minority of the population holds more than one tertiary degree. In very few cases we encounter the problem that a respondent provides information on two different degrees in one survey year. This only happens in years when respondents fill in the person as well as the biography questionnaire. In these cases we prioritize the information from the person questionnaire as it refers to the current situation while the biography questionnaire contains retrospective information. Furthermore, there are cases who report an applied university degree and a university degree in the biography questionnaire. In these cases, the variables contain information on the university degree only.
The variable is coded according to a slightly collapsed version of the classification on types of tertiary degrees ("Prüfungsgruppen und Abschlussprüfungen") provided by the Statistisches Bundesamt (2009). Since 2010 the data were coded according to the classification presented here. In the variable PGDEGREE we recoded the original values from years 2009 and earlier. As the newer version is more precise this could be done with hardly any loss of information. Some categories are collapsed. Category 16 was mostly likely coded as 15 in earlier years, 34 as 35 and 43 as 44. The original values of the data collected from the person questionnaire of 2009 are stored in the respective variables in the dataset P.

Because detailed information on the type of tertiary degree is not assessed in the Questionnaire of the SOEP Innovation Sample, PGDEGREE is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGDEGREE is coded to "-5" (not contained in questionnaire).
Stat. Bundesamt (2009): Bildung und Kultur. Studierende an Hochschulen, Fachserie 11 Reihe 4.1, Wiesbaden: 449ff, Übersicht 2: „Prüfungsgruppen und Abschlussprüfungen".

pgtraina – Apprenticeship, two-digit occupation KldB92

1	[1] Agricultural Occupations (Crops)	21
2	[2] Agricultural Occupations (Livestock)	15
3	[3] Administrative/Advisory/Technical Specialist In Agriculture	1
5	[5] Horticultural Occupations	19
6	[6] Forestry and Hunting Occupations	0
7	[7] Mineworkers	3
8	[8] Mineral Exploitation and Processing	1
10	[10] Stonemasons	0
11	[11] Manufacturers of Construction Materials	0
12	[12] Ceramicists	0
13	[13] Glass Manufacturing Occupations	1
14	[14] Chemical Industry Occupations	4
15	[15] Plastics Manufacturing Occupations	1
16	[16] Paper Manufacturing and Processing	3
17	[17] Printing Occupations	14
...	(65 rows omitted)	1516
89	[89] Pastoral Occupations	0
90	[90] Personal Care Occupations	46
91	[91] Occupations in Hotels and Hospitality	15
92	[92] Occupations in Domestic and Nutritional Science	16
93	[93] Cleaning and Waste Management Occupations	0
96	[96] Other	0
97	[97] Family members providing assistance,not in agriculture,not otherw. mntnd	0
98	[98] Workers, (still) without specific occupation	0
99	[99] Workers, responsibilities not specified	11
-1	[-1] No Answer	13
-2	[-2] Does Not Apply	13625
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	46965
-6	[-6] Questionnaire Version With Modified Filtering	0

The variable is designed to provide information on the occupation of vocational training which adds details to the information recorded in the variable PGBBIL01. In addition to the variable PGTRAINA we provide the variables PGTRAINB, PGTRAINC and PGTRAIND. All these variables record the occupation of vocational training. The difference is that PGTRAINA contains information on vocational training within the German dual system which combines firm-based and school-based training (apprenticeship). PGTRAINB is designed to provide information on the occupation of full-time school based vocational training. PGTRAINC contains information on level vocational training (e.g., Meister, Techniker). PGTRAIND is designed to provide information on the occupation of civil servant training ("Beamtenausbil-

dung"). We describe in brief detail the construction of the variable PGTRAINA. PGTRAINB, PGTRAINC and PGTRAIND are constructed in an analogous manner.
The data of the generated variable PGTRAINA stem from two sources: 1. Person questionnaire: Each year since 1985 respondents are asked if they have left education since the beginning of the year prior to the survey and which degrees they have obtained. This part of the questionnaire contains an open question on the type and the field of newly obtained tertiary degrees. This information is coded and used for the generation of the variables PGTRAINA. 2. Biography questionnaire: Similar information is collected from respondents who fill in the biography questionnaire (usually during the first two years of participation in the panel). In contrast to the information from the person questionnaire the questions do not refer to currently obtained vocational qualifications but to qualifications obtained during the time before being part of the SOEP sample.
In the variable PGTRAINA we combine these two types of information. Each year the variable contains the most recently collected information.
If you want to take into account that a person holds two vocational qualifications you have to combine the information from all available years. In few cases we encounter the problem that a respondent provides information on two different apprenticeships in one survey year. This only happens once, namely in years when respondents fill in the person as well as the biography questionnaire. In these cases we prioritize the information from the person questionnaire as it refers to the current situation while the biography questionnaire contains retrospective information.
Because detailed information on the specific fields of occupation is not assessed in the Questionnaire of the SOEP Innovation Sample, PGTRAINA is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGTRAINA is coded to "-5" (not contained in questionnaire).
Hartmann/Schütz (2002): Die Klassifikation der Berufe und der Wirtschaftzweige im Sozio-oekonomischen Panel – Neuvercodung der Daten 1984 – 2001. Infratest Sozialforschung, München.

pgtrainb – Vocational school, two-digit occupation KldB92

1	[1] Agricultural Occupations (Crops)	4
2	[2] Agricultural Occupations (Livestock)	0
3	[3] Administrative/Advisory/Technical Specialist In Agriculture	0
5	[5] Horticultural Occupations	0
6	[6] Forestry and Hunting Occupations	1
7	[7] Mineworkers	0
8	[8] Mineral Exploitation and Processing	0
10	[10] Stonemasons	0
11	[11] Manufacturers of Construction Materials	0
12	[12] Ceramicists	0
13	[13] Glass Manufacturing Occupations	0
14	[14] Chemical Industry Occupations	0
15	[15] Plastics Manufacturing Occupations	0
16	[16] Paper Manufacturing and Processing	0
17	[17] Printing Occupations	1
...	(65 rows omitted)	248
89	[89] Pastoral Occupations	0
90	[90] Personal Care Occupations	2
91	[91] Occupations in Hotels and Hospitality	1

92	[92] Occupations in Domestic and Nutritional Science	7
93	[93] Cleaning and Waste Management Occupations	0
96	[96] Other	0
97	[97] Family members providing assistance,not in agriculture,not otherw. mntnd	0
98	[98] Workers, (still) without specific occupation	0
99	[99] Workers, responsibilities not specified	0
-1	[-1] No Answer	1
-2	[-2] Does Not Apply	14334
-3	[-3] Answer Improbable	2
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	47689
-6	[-6] Questionnaire Version With Modified Filtering	0

The variable is designed to provide information on the occupation of full-time school based vocational training (e.g., Berufsfachschule, Schule des Gesundheitswesens, Handelsschule). See the description of variable PGTRAINA for more details on the construction and the values of the variable.
Because detailed information on the specific fields of occupation is not assessed in the Questionnaire of the SOEP Innovation Sample, PGTRAINB is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGTRAINB is coded to "-5" (not contained in questionnaire).

pgtrainc – Higher vocational school, two-digit occupation KldB92

1	[1] Agricultural Occupations (Crops)	3
2	[2] Agricultural Occupations (Livestock)	2
3	[3] Administrative/Advisory/Technical Specialist In Agriculture	3
5	[5] Horticultural Occupations	1
6	[6] Forestry and Hunting Occupations	1
7	[7] Mineworkers	0
8	[8] Mineral Exploitation and Processing	0
10	[10] Stonemasons	0
11	[11] Manufacturers of Construction Materials	0
12	[12] Ceramicists	0
13	[13] Glass Manufacturing Occupations	0
14	[14] Chemical Industry Occupations	0
15	[15] Plastics Manufacturing Occupations	0
16	[16] Paper Manufacturing and Processing	0
17	[17] Printing Occupations	1
...	(65 rows omitted)	159
89	[89] Pastoral Occupations	0
90	[90] Personal Care Occupations	14
91	[91] Occupations in Hotels and Hospitality	0
92	[92] Occupations in Domestic and Nutritional Science	2
93	[93] Cleaning and Waste Management Occupations	0
96	[96] Others	0
97	[97] Family members providing assistance,not in agriculture,not otherw. mntnd	0
98	[98] Workers, (still) without specific occupation	0
99	[99] Workers, responsibilities not specified	0
-1	[-1] No Answer	0

-2	[-2] Does Not Apply	14415
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	47689
-6	[-6] Questionnaire Version With Modified Filtering	0

The variable is designed to provide information on the occupation of higher level vocational training (e.g., Meister, Techniker). See the description of variable PGTRAINA for more details on the construction and the values of the variable.
Because detailed information on the specific fields of occupation is not assessed in the Questionnaire of the SOEP Innovation Sample, PGTRAINC is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGTRAINC is coded to "-5" (not contained in questionnaire).

pgtraind – Civil servant training, two-digit occupation KldB92

1	[1] Agricultural Occupations (Crops)	0
2	[2] Agricultural Occupations (Livestock)	0
3	[3] Administrative/Advisory/Technical Specialist In Agriculture	0
5	[5] Horticultural Occupations	0
6	[6] Forestry and Hunting Occupations	2
7	[7] Mineworkers	0
8	[8] Mineral Exploitation and Processing	0
10	[10] Stonemasons	0
11	[11] Manufacturers of Construction Materials	0
12	[12] Ceramicists	0
13	[13] Glass Manufacturing Occupations	0
14	[14] Chemical Industry Occupations	0
15	[15] Plastics Manufacturing Occupations	0
16	[16] Paper Manufacturing and Processing	0
17	[17] Printing Occupations	0
...	(65 rows omitted)	97
89	[89] Pastoral Occupations	0
90	[90] Personal Care Occupations	0
91	[91] Occupations in Hotels and Hospitality	0
92	[92] Occupations in Domestic and Nutritional Science	0
93	[93] Cleaning and Waste Management Occupations	0
96	[96] Other	0
97	[97] Family members providing assistance,not in agriculture,not otherw. mntnd	0
98	[98] Workers, (still) without specific occupation	0
99	[99] Workers, responsibilities not specified	0
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	11449
-3	[-3] Answer Improbable	2
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	50740
-6	[-6] Questionnaire Version With Modified Filtering	0

The variable is designed to provide information on the occupation of civil servant training ("Beamtenausbildung"). See the description of variable PGTRAINA for more details on the

construction and the values of the variable.
Because detailed information on the specific fields of occupation is not assessed in the Questionnaire of the SOEP Innovation Sample, PGTRAIND is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGTRAIND is coded to "-5" (not contained in questionnaire).

pgfdt_f – Data source FIELD, DEGREE, TRAIN

1	[1] Individual Questionnaire	915
2	[2] Gap Questionnaire (temporary drop-outs)	0
3	[3] Biographical Questionnaire	1778
4	[4] Various Sources	26
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	12606
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	46965
-6	[-6] Questionnaire Version With Modified Filtering	0

Because detailed information on PGFIELD, PGDEGREE and PGTRAIN is not assessed in the Questionnaire of the SOEP Innovation Sample, PGFDT_F is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGFDT_F is coded to "-5" (not contained in questionnaire).

pgbilztch – Change in Education since last survey / last year

0	[0] Consistent educational information since last survey	10331
1	[1] Inconsistent educational information since last survey	322
2	[2] Inconsistent educational information since last year	441
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	3507
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	47689
-6	[-6] Questionnaire Version With Modified Filtering	0

Because detailed information on PGFIELD, PGDEGREE and PGTRAIN is not assessed in the Questionnaire of the SOEP Innovation Sample, PGBILZTCH is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGBILZTCH is coded to "-5" (not contained in questionnaire).

pgbilztev – Change in Education, total observed period

0	[0] Consistent educational information	9346
1	[1] Inconsistent educational decline	2015
2	[2] Inconsistent educational increase	2553
3	[3] Inconsistent educational decline and increase	56
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	1355

-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	46965
-6	[-6] Questionnaire Version With Modified Filtering	0

Because detailed information on PGFIELD, PGDEGREE and PGTRAIN is not assessed in the Questionnaire of the SOEP Innovation Sample, PGBILZTEV is not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGBILZTEV is coded to "-5" (not contained in questionnaire).

pgsndjob – Current gross secondary income in euros

? How much was your gross income respectively your allowance gained from all side jobs in the last month? // EUR *(from: soep-is/soep-is-2019/Q505:pnebbr[28513])*

0		1844
1		3
2		1
5		5
9		1
10		21
12		3
13		1
15		10
16		2
18		2
20		45
21		1
22		1
24		2
...	(334 rows omitted)	3393
4000		2
4100		1
4395		1
5000		2
6000		5
6800		1
7000		2
10000		1
15000		1
-1	[-1] No Answer	741
-2	[-2] Does Not Apply	56198
-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 variable PGSNDJOB represents the imputed current gross labor income from second job generated for all SOEP respondents in each respective wave. Because missing information on PGSNDJOB is not imputed in the SOEP Innovation Sample, PGSNDJOB not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012).

For this purpose, PGSNDJOB is coded to "-5" (not contained in questionnaire).

pgimpsnd – Imputation flag for SNDJOB

0	[0] not imputed	758
1	[1] Imputed	114
-1	[-1] No Answer	0
-2	[-2] Does Not Apply	16154
-3	[-3] Answer Improbable	0
-4	[-4] Inadmissible Multiple Answer	0
-5	[-5] Not Contained In Questionnaire	45264
-6	[-6] Questionnaire Version With Modified Filtering	0

The variable PGIMPSND indicates imputations of item nonresponse in the variable PGSNDJOB (current gross labor income from second job). Because missing information on PGSNDJOB is not imputed in the SOEP Innovation Sample, PGIMPSND not generated for Sample E (since 2012), I (since 2011) and the supplementary samples (since 2012). For this purpose, PGIMPSND is coded to "-5" (not contained in questionnaire).