SOEP The German Socio-Economic Panel study

483

SOEP Survey Papers

Series D – Variable Descriptions and Coding

SOEP – The German Socio-Economic Panel study at DIW Berlin 2018

SOEP-Core v33.1 – Documentation of Person-related Status and Generated Variables in $PGEN

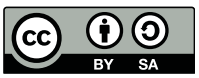
SOEP Group

DIW SOEP

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

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

Please cite this paper as follows:
SOEP Group, 2018. SOEP-Core v33.1 – Documentation of Person-related Status and Generated Variables in $PGEN. SOEP Survey Papers 483: 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 study at DIW Berlin

SOEP-Core v33.1 – Documentation of Person-related Status and Generated Variables in $PGEN

SOEP Group

2018

The files $PGEN are part of a collection, which is released with doi:10.5684/soep.v33.1.

Contents

1 General Information 4

2 Identifiers 4
persnr – Never Changing Person ID . . . 4
hhnr – Original Household Number . . . 4
hhnrakt – hhnrakt Current Wave HH Number (=$hhnr) [generic] . . . 4
$hhnr – Current Wave HH Number (=HHNRAKT) [generic] . . . 4

3 Own Nationality and Residential Status 4
nation$$ – Nationality [generic] . . . 4
status_asyl – Status quo, Application for Asylum . . . 5
status_refu – Refugee residence status . . . 5

4 Familiy Status and Partnership 6
partz$$ – Partner Indicator [generic] . . . 6
partnr$$ – Partner Person Number [generic] . . . 6
$famstd – Marital Status In Survey Year [generic] . . . 7

5 Wages and Salary 8
labgro$$ – Current Gross Labor Income in Euro [generic] . . . 8
impgro$$ – Imputation flag for LABGROxx [generic] . . . 8
labnet$$ – Current Net Labor Income in Euro [generic] . . . 8
impnet$$ – Imputation flag for LABNETxx [generic] . . . 9
sndjob$$ – Current Gross Secondary Income In Euro [generic] . . . 9
impsnd$$ – Imputation Flag For SNDJOBxx [generic] . . . 10

6 Current Employment Status 10
stib$$ – Occupational Position [generic] . . . 10
emplst$$ – Employment Status [generic] . . . 11
lfs$$ – Labor Force Status [generic] . . . 12
jobch$$ – Occupational Change [generic] . . . 13
autono$$ – Autonomy In Occupational Actions [generic] . . . 15

7 Current Occupation 16
isco88_$$ – Current Occupational Classification (ISCO-88 Com) [generic] . . . 16
isco08_$$ – Current Occupational Classification (ISCO-08) [generic] . . . 17
isei88_$$ – Last Reached Isei Value (International Socio-Economic Index) [generic] . . . 18
kldb92_$$ – Current Occupational Classification (KldB92) [generic] . . . 19
kldb2010_$$ – Current Occupational Classification (KldB2010) [generic] . . . 20
mps92_$$ – Last Reached Mps Value (Magnitude-Prestige-Skale, Wegener) [generic] . . . 20
siops88_$$ – Last Reached Siops Value (Std. Internat. Occupational Prestige Scale) [generic] 21
egp88_$$ – Last Reached Egp Value (Erikson, Goldthorpe, Portocarero) [generic] . . . 22
erljob$$ – Working In Occupation Trained For [generic] . . . 23
ausb$$ – Required Training For Job [generic] . . . 23
$erwzeit – Length Of Time With Firm [generic] . . . 24
$tatzeit – Actual Work Time Per Week [generic] . . . 25
$vebzeit – Agreed Upon Work Time Per Week [generic] . . . 25
$uebstd – Overtime Per Week [generic] . . . 26
oeffd$$ – Civil Service [generic] . . . 26
nace$$ – Industry Occupation [pbra] (NACE Rev. 1.1, Sector) [generic] . . . 26

nace2$$ – Industry Occupation [pbra] (NACE Rev. 2, Sector) [generic] . . . 27
betr$$ – Size of the Company [generic] . . . 28
allbet$$ – Core Category Size Of The Company [generic] . . . 29

8 Last Occupation **30**
jobend$$ – Reasons for occupational change [generic] . . . 30

9 Employment History **31**
expft$$ – Working Experience Full-Time Employment [generic] . . . 31
exppt$$ – Working Experience Part-Time Employment [generic] . . . 32
expue$$ – Unemployment Experience [generic] . . . 33

10 School, Higher and Vocational Education **34**
isced97_$$ – ISCED-1997-Classification [generic] . . . 34
isced11_$$ – ISCED-2011-Classification [generic] . . . 35
casmin$$ – CASMIN Classification [generic] . . . 35
$bilzeit – Amount Of Education Or Training In Years [generic] . . . 36
$psbil – School-Leaving Degree [generic] . . . 36
$pbbil01 – Vocational Degree Received [generic] . . . 37
$pbbil02 – College Degree [generic] . . . 37
$pbbil03 – No Vocational Degree [generic] . . . 38
$psbilo – School-Leaving Degree East Germany [generic] . . . 39
$pbbilo – Vocational Degree Received East Germany [generic] . . . 39
$psbila – School-Leaving Degree Outside Germany [generic] . . . 40
$pbbila – Vocational Degree Outside Germany [generic] . . . 40
field$$ – Field of tertiary education [generic] . . . 41
degree$$ – Type of tertiary degree [generic] . . . 42
traina$$ – Apprenticeship - two-digit occupation KldB92 [generic] . . . 44
trainb$$ – Vocational school - two-digit occupation KldB92 [generic] . . . 46
trainc$$ – Higher vocational school - two-digit occupation KldB92 [generic] . . . 47
traind$$ – Civil servant training - two-digit occupation KldB92 [generic] . . . 48
fdt_f$$ – Data source FIELD, DEGREE, TRAIN [generic] . . . 49
bilztch$$ – Change in Education since last survey / last year [generic] . . . 49
bilztev$$ – Change in Education, total observed period [generic] . . . 50

11 Political and Religious Attitudes **50**
picoriga16 – Party Identification in Country of Origin . . . 50
picorigb16 – Party Identified with in Country of Origin . . . 51
picorigc16 – Party Family of PI in Country of Origin . . . 51
bgp_reli – Belong To Church Or Religious Community . . . 52

12 Information on the Interview **52**
month$$ – Month Of Interview [generic] . . . 52
mode$$ – Interview Method [generic] . . . 53

1 General Information

The $PGEN-files contain user friendly data on the individual level which are consolidated from different sources. The plausibility is in many respects longitudinally validated, therefore the data here are in most situations superior compared to the data in $P.
The file contains one row for each person (persnr is unique) with a completed personal or youth questionnaire. These are the persons where in PPFAD $netto has the values 10–17 or 19 which is equivalent for values 1 and 5 in $netold.
While frequencies are calculated from the most recent $PGEN the here presented information is basically valid for all $PGEN files. Therefore many variable names depicted here have the generic form $name or name$$ and are flagged with [generic].
From version v32 on ERWTYP is not a part of data delivery. The information related to ERWTYP categories can be found in the variable(s) JOBCH$$, EMPLST$$, LFS$$.
The long version of the $pgen files is file pgen, which contains all $pgen information of all waves. The variable names in the long file got the prefix pg and the suffixes were deleted.

2 Identifiers

persnr – Never Changing Person ID

The central individual identifier across time is PERSNR, which is fixed over time (and of course datasets).

hhnr – Original Household Number

The identifier of the household, when it is sampled and selected for interviewing for the first time. The hhnr is attached to all persons living in this household and all new persons inherit this identifier, when they are born or move in a SOEP household. It is fixed no matter how often a person changes the household in the course of time.

hhnrakt – hhnrakt Current Wave HH Number (=$hhnr) [generic]

This identifier groups all persons into households at the time of the most recent wave. [This information can be related to a specific variable and is not necessarily generic.]

$hhnr – Current Wave HH Number (=HHNRAKT) [generic]

This identifier groups all persons into households at the time of the most recent wave. [This information can be related to a specific variable and is not necessarily generic.]

3 Own Nationality and Residential Status

nation$$ – Nationality [generic]

1	[1] Germany	21985
2	[2] Turkey	408
3	[3] Ex-Yugoslavia	7
4	[4] Greece	184
5	[5] Italy	260
6	[6] Spain	91
7	[7] Ex-GDR (Country Of Origin Only)	0

10	[10] Austria	64
11	[11] France	36
12	[12] Benelux	0
13	[13] Denmark	11
14	[14] Great Britain	28
15	[15] Sweden	7
16	[16] Norway	5
17	[17] Finland	8
...	(159 rows omitted)	6816
177	[177] Bhutan	0
178	[178] Rwanda	2
179	[179] Malawi	0
180	[180] Bessarabia	0
183	[183] Niger	4
188	[188] Chechnya	13
222	[222] Eastern Europe	0
333	[333] Other Unspecified Foreign Country	0
999	[999] ethnic minorities (e.g. Yazidi, Roma)	2
-1	[-1] No Answer	0
-2	[-2] Does not apply	4
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
This variable is designed to integrate the information on respondent's nationality for all subsamples. Since some members of Sample B (persons with Turkish, Italian, Spanish, Greek, and Yugoslavian citizenship) received the question items in their own language up to 1995, to carry out an integrated analysis with Sample B, the user must obtain this information from the corresponding $PAUSL files and add it to the individual data. The variable NATION$$ thus offers a variable on nationality for all subsamples. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

status_asyl – Status quo, Application for Asylum

1	[1] Protection Status (Recognized Refugee, Entitled to Asylum, other Protection)	2482
2	[2] Temporary Suspension of Deportation	163
3	[3] Request to Leave Germany	76
4	[4] Decision Regarding Application for Asylum Still open	1516
-1	[-1] No Answer	50
-2	[-2] Does not apply	25426
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	222
-6	[-6] Version of questionnaire with modified filtering	0

status_refu – Refugee residence status

1	[1] In Proceedings	1446
2	[2] Protection Status (Asylum, Refugee Status, Settlement Permit)	2469
3	[3] Temporary Suspension of Deportation	300
4	[4] Other	203
-1	[-1] No Answer	109
-2	[-2] Does not apply	25186
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	222
-6	[-6] Version of questionnaire with modified filtering	0

4 Familiy Status and Partnership

partz$$ – Partner Indicator [generic]

0	[0] No partner	10128
1	[1] Spouse, registered partner	16895
2	[2] Partner	2517
3	[3] Probably spouse, registered partner	125
4	[4] Probably partner	111
-1	[-1] No Answer	159
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
The variable PARTZ$$ generated in the context of the partner identifier (PARTNR$$) to describe whether a person in a SOEP 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 assigned to all single persons living in households and those with partners outside the household. Codes 1 to 4 describe relationships. To assign Codes 1 and 2, the partnership has to be definable from the perspective of both partners unanimously. If conflicting information exists between partners, the codes 3 or 4 are assigned. If it is unclear whether an individual has no partner or whether she forms a couple with one other household member, we assign the code -1. Registered partnerships (civil unions) for same-sex couples were introduced in Germany in 2001. Though, registered partnerships are legally not equal to marriage, they are listed in the same category. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Martin Kroh (Tel. +49-30-89789-678)

partnr$$ – Partner Person Number [generic]

Waves: all
Partner indicators have the purpose of defining couples in SOEP households and thus to make possible analyses on the dyadic level. Persons without spouse and (cohabitating) partner receive a missing code "-2" (=does not apply). Also, the variable PARTZ$$ is coded -1, 0, 3, 4 in these cases. In couples, PARTNR$$ is the value of the unchanging person ID

number (=PERSNR) of the partner. The assignment of the partner ID within households is based on four sources of information: A question in the person-file, that asks (unmarried) respondents to identify their partner in the household (bgppnr in 2016), the household matrix reported by the head of household at the beginning of the interview (bgstell in 2016), the partnership biography in the lifehistory calendar reported by new respondents (see also, biomars), and self-reports on marital status and life events, such as marriage, move in with partner, separation, etc (lela). In unclear cases, due to temporal non-response for instance, we also consider longitudinal information from previous and prospective waves. Moreover, PARTNR$$ is self-consistent between two individuals. For analyses of partner relationships, this information can be used to link all persons with their respective partners, and all information on both partners can also be stored in a common dataset. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Martin Kroh (Tel. +49-30-89789-678)

$famstd – Marital Status In Survey Year [generic]

1	[1] Married	16807
2	[2] Married, But Separated	673
3	[3] Single	8107
4	[4] Divorced	2305
5	[5] Widowed	1431
6	[6] husband/wife abroad	380
7	[7] Registered Same-Sex Partnership, Living Together	55
8	[8] Registered Same-Sex Partnership, Living Apart	12
-1	[-1] No Answer	60
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	105
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: 1984-2012
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. Information on marital status when a child was born (provided in the biography information) is not used here, so contradicting information to BIOBIRTH might still be possible. 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. Data on marital status provided here is generated from and hence consistent with the corresponding BIOMARSY data, although value labels differ between $FAMSTD and SPELLTYP. Note that marital status in a specific wave can be different between data distributions due to consistency checks using up-to-date information from following waves. Because of the newly introduced construction process of BIOMARSY, changes might be more than usual. Above, the partner indicator PARTZ$$ supplied in the $PGEN data files as well might not match the information provided in $FAMSTD in its en-

tirety. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Paul Schmelzer (Tel. +49-30-89789-526, pschmelzer@diw.de)

5 Wages and Salary

labgro$$ – Current Gross Labor Income in Euro [generic]

Waves: all
The variable LABGRO$$ represents the imputed current gross labor income generated for all SOEP respondents who are employed in a main job in each respective wave. Income details are consistently provided in euros for all waves. Item nonresponse is imputed in a two-stage procedure: first, with the "Row-and-Column" method of Little und Su (1989) using individual longitudinal data as well as cross-sectional trend data (cf. Joachim R. Frick and Markus M. Grabka (2005): Item-Non-Response on Income Questions in Panel surveys: Incidence, Imputation and the Impact on the Income Distribution. Allgemeines Statistisches Archiv (ASTA) 89, 49-61). Alternatively, if no individual longitudinal information is available, we base the imputation on a regression using different Mincer covariates, also taking into account current net labor income. If both types of income information are lacking, first we impute current net labor income and then current gross labor income. Imputed values are flagged (IMPGRO$$).
The original variables coming from the $P-files and are: ap3301, bp4301, cp5201, dp4401, ep4401, fp4501, gp4301, hp5401, ip5401, jp5401, kp6401, lp5301, mp4701, np5401, op4501, pp6001, qp5601, rp5701, sp5801, tp7601, up5901, vp7101, wp5901, xp7301, yp6801, zp7201, bap6101, bbp7301, bcp5901, bdp7701, bep5701, bfp10201, bgp8101. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

impgro$$ – Imputation flag for LABGROxx [generic]

0	[0] Not imputed	14487
1	[1] Imputed	1314
-1	[-1] No Answer	0
-2	[-2] Does not apply	13377
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	757
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
The variable IMPGRO$$ designates imputations of item non-response in the variable LABGRO$$ (current gross labor income). IMPGRO$$ can take the value 0 = "no imputation", 1 = "imputed income statement" and - 2 = "does not apply, not working". [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

labnet$$ – Current Net Labor Income in Euro [generic]

Waves: all
The variable LABNET$$ represents the generated and imputed current net labor income in a main job of all persons in SOEP working in the respective wave. 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 (IMPNET$$).
The original variables coming from the $P-files and are: ap3302, bp4302, cp5202, dp4402, ep4402, fp4502, gp4302, hp5402, ip5402, jp5402, kp6402, lp5302, mp4702, np5402, op4502, pp6002, qp5602, rp5702, sp5802, tp7602, up5902, vp7102, wp5902, xp7302, yp6802, zp7202, bap6102, bbp7302, bcp5902, bdp7702, bep5702, bfp10202, bgp8102. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

impnet$$ – Imputation flag for LABNETxx [generic]

0	[0] Not imputed	14800
1	[1] Imputed	1001
-1	[-1] No Answer	0
-2	[-2] Does not apply	13377
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	757
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
The variable IMPNET$$ designates imputations of item non-response in the variable LABNET$$ (current net labor income). IMPNET$$ can take the value 0 = "no imputation", 1 = "imputed income statement" and - 2 = "does not apply, not working". [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

sndjob$$ – Current Gross Secondary Income In Euro [generic]

Waves: all
The variable SNDJOB$$ represents the imputed current gross labor income from second job generated for all SOEP respondents in each respective wave. Income details are consistently provided in euros for all waves. Item nonresponse is imputed 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 a subset of different Mincer covariates. Information about gross income from second job is firstly asked in wave 1995 (L). In 2013 information about income from second jobs was not collected for sub-sample M. We use information from the the subsequent survey year (2013) about income from second job in the previous year (variable BEP2C03). If

persons from the sub-sample M stated that he/she had income in the previous year from second jobs, then IMPSND13 was set to -1 and SNDJOB$$ had been imputed. Imputed values are flagged (IMPSND$$).
The original variables coming from the $P-files and are: lp7702, mp5802, np5802, op4902, pp6602, qp6302, rp6602, sp6602, tp8402, up67, vp79, wp70, xp81, yp76, zp78, bap69, bbp79, bcp67, bdp85, bep64, bfp113, bgp97. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

impsnd$$ – Imputation Flag For SNDJOBxx [generic]

0	1469
1	156
-2	27553
-5	757

Waves: all
The variable IMPSND$$ designates imputations of itemnonresponse in the variable SNDJOB$$ (current gross labor income from second job). IMPSND$$ can take the value 0 = “no imputation”, 1 = “imputed income statement” and - 2 = “does not apply, not working”. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

6 Current Employment Status

stib$$ – Occupational Position [generic]

0	[0] Do Not Know	0
10	[10] Not Employed	5446
11	[11] In Education	1166
12	[12] Unemployed, Not Employer	1428
13	[13] Pensioner	5301
15	[15] Military, Community Service	57
110	[110] Apprentice	0
120	[120] Apprentice,Trainee Industry Technology	446
130	[130] Apprentice, Trainee Trade And Commerce	218
140	[140] Trainee, Intern	164
150	[150] Aspirant	0
210	[210] Untrained Worker	761
220	[220] Semi-Trained Worker	1271
230	[230] Trained Worker	1189
240	[240] Foreman, Team Leader	175
...	(21 rows omitted)	2725
522	[522] Trained Employee With Simple Tasks	1355
530	[530] Qualified Professional	3945
540	[540] H. Qualified Professional	2239
550	[550] Managerial	191
610	[610] Low-Level Civil Service	23
620	[620] Middle-Level Civil Service	189

630	[630] High-Level Civil Service	454
640	[640] Executive Civil Service	263
999	[999] Employed Without StiB Info	0
-1	[-1] No Answer	308
-2	[-2] Does not apply	621
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
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):

value	label
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

The categories (150) and (310) to (340) were only assigned to respondents in East Germany in 1990. In STIB$$, 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. Moreover, if there is missing information on pension receipt, additional information from ARTKALEN (retrospective information from the activity calendar for the previous year) is used in the generation process to determine if a person was in retirement or early retirement (Vorruhestand) at the time of the interview.
The code (-2) is assigned to first time respondents aged 16 or 17 who answer the youth questionnaire (since wave W (2006)). Furthermore, the code (-2) is assigned to the group of 229 refugee teenagers in 2016. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

emplst$$ – Employment Status [generic]

1	[1] Full-Time Employment	9384
2	[2] Regular Part-Time Employment	3883
3	[3] Vocational Training	857
4	[4] Marginal, Irregular Part-Time Employment	1729
5	[5] Not Employed	13826

6	[6] Sheltered workshop	33
-1	[-1] No Answer	1
-2	[-2] Does not apply	222
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
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.
Since the beginning of the SOEP in the year 1984, a consistent status variable has been used to differentiate among different types of employment status. 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. From 1998 on, the additional category "sheltered workshop" is included for disabled persons in sheltered employment.
EMPLST$$ supplements the variable LFS$$, which differentiates among persons who are not employed. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

lfs$$ – Labor Force Status [generic]

1	[1] Non-Working	5883
2	[2] NW-Age 65 And Older	4389
3	[3] NW-In Education-Training	1284
4	[4] NW-Maternity Leave	469
5	[5] NW-Military-Community Service	13
6	[6] NW-Unemployed	1277
8	[8] NW-But Sometimes Sec. Job	235
9	[9] NW-but work past 7 days	208
10	[10] NW-But Reg. Sec. Job	233
11	[11] Working	15698
12	[12] Working But NW Past 7 Days	246
-1	[-1] No Answer	0
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
This variable is based on the annual question on current employment status, combined with additional information on activities of non-working individuals. The number of values assigned has been based, since the beginning of the SOEP in the year 1984, on a large number of highly differentiated answer categories. It is designed to provide consistent longitudinal data on labor force participation across all waves.

LFS$$ 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):

value	label
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

LFS$$ supplements the variable EMPLST$$, which differentiates among persons who are employed. In 2016, the questionnaire for refugees contains a new labor market status category „vocational internship / betriebliches Praktikum", which is integrated in the category 10 „education" of the LFS$$ variable. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

jobch$$ – Occupational Change [generic]

1	[1] Not Employed	13775
2	[2] Employed No Change	11892
3	[3] Employed No Info If Change	309
4	[4] Employed With Change	3432
5	[5] First Time Employed	304
-1	[-1] No Answer	223
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
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.
JOBCH$$ is generated based on the central filter variable, which indicates whether a respon-

dent has changed jobs since the beginning of the previous year. A job change can be within one firm as well as a change to another firm. 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, JOBCH$$ indicates whether a respondent has changed jobs since the beginning of the previous year. The variable is calculated for all waves, and the codes are assigned independently of the respondent being a first-time or follow-up respondent.
The variable is also designed to identify respondents who have entered employment for the first time. Up to 1993, first-time respondents did not answer the question about job change. Therefore, for first-time respondents up to 1993, the variable was generated by using the information on the start date with the current employer and the respondent's age at entrance into his/her first job.
The variable is also designed to provide consistent longitudinal information on job changes. The JOBCH$$ 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".

In 2013 the respondents of the newly introduced migration sample (M1) were not asked whether they have changed jobs since the beginning of the previous year, therefore the generation of JOBCH$$ for the migration sample was modified in 2013:

1. Respondents who are not employed were coded (1).
2. Respondents who are still in the same occupation and position and are working for the same employer as they had worked in their first job in Germany were coded (2).
3. Respondents who have entered the firm they are currently working after the 31th of December 2011 were coded (4).
4. If a respondent is in her first vocational training this was coded as (5).
5. Respondents who are employed but for whom no further information could be used were coded (3). In 2014 there was again a uniform questionnaire for all respondents.

In 2015 the respondents of the newly introduced migration sample (M2) were not asked whether they have changed jobs since the beginning of the previous year. Furthermore respondent within migration sample M1 are decomposed in first-time respondents with and follow-up respondents without question about job change in the previous year embodied by the questionnaire. such that generation of JOBCH$$ for the migration sample in 2015 was following:

1. For follow-up respondents generation as in case of a uniform questionnaire for all respondents to be pursued.

2. In the case of first-time respondents in M1 sample and for all respondents from M2 sample the rule of thumb is as in 2013.

In 2016 the respondents of the newly introduced refugee samples (M3 and M4, psample==17) were not asked whether they have changed jobs since the beginning of the previous year. The generation of JOBCH$$ for these samples corresponds to the rule of thumb is as for the forst-time respondent in the M1 sample with one modification. In step 4, the code (5) was assigned to respondents who are currently in a vocational training, were not occupied prior to their arrival in Germany, had neither a vocational training, nor a university degree and are in their first occupational relationship in Germany. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

autono$$ – Autonomy In Occupational Actions [generic]

0	[0] Job: Autonomy (low=1, high=5; 0=apprentice)	912
1	[1] Job: Autonomy (low=1, high=5; 0=apprentice)	2025
2	[2] Job: Autonomy (low=1, high=5; 0=apprentice)	3742
3	[3] Job: Autonomy (low=1, high=5; 0=apprentice)	5135
4	[4] Job: Autonomy (low=1, high=5; 0=apprentice)	3211
5	[5] Job: Autonomy (low=1, high=5; 0=apprentice)	552
-1	[-1] No Answer	590
-2	[-2] Does not apply	13768
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
This variable gives the occupational autonomy for all employed persons. It offers an alternative to the ISCO-based scales on occupational status (ISEI$$), class (EGP$$), or prestige (SIOPS$$). AUTONO$$ is the simplest variable based on the scales of "occupational position" in terms of its construction, and strongly correlated with the Treiman Prestige Scale (SIOPS$$).
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. [This information can be related to a specific variable and is not necessarily generic.]
Detailed description: Hoffmeyer-Zlotnik, Jürgen H.P., and Alfons J. Geis (2003) Berufsklassifikation und Messung des beruflichen Status/ Prestige. In: ZUMA-Nachrichten 52, Jg. 27, Mai 2003. pp. 125-138.
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

7 Current Occupation

Not all employed persons are asked the question about occupation on an annual basis. In years with a partial survey – 1985, 1986, 1987, 1988, 1990 (West), 1992 (West), 1994, 1996, 1999, 2001, 2003, 2005, 2006, 2008, 2010, 2012, 2014, and 2016 – only those employed persons who changed jobs and first-time respondents are asked to provide up-to-date information. Therefore, in years with a partial survey or in case of non-response the variables in this section usually contain available information from the previous year (persons with JOBCH$$-category (2) "employed, no change"). For some persons without a job change who updated the information on their current occupation without being asked, up-to-date information is used. The scores which are derived from the occupational scores contain information on the last attained value.
The missing values in variables of codes for economic activities or occupations and derived scores (NACE, ISCO, KldB, ISEI, SIOPS, EGP, MPG) should be interpreted as follows. -1: there was a response, but no code could be assigned or no score could be derived, -2: there was no response which could have been coded and -8 means this type of code or score is not available for this year.
The information on the current occupation is not neccessarily consistent to the current employment status, because they are based on different sources of information

isco88_$$ – Current Occupational Classification (ISCO-88 Com) [generic]

0	[0] Soldiers	0
100	[100] Soldiers	38
1000	[1000] Legislators, Senior Officials and Managers	0
1100	[1100] Legislators, Senior Officials and Managers	0
1110	[1110] Legislators, Senior Officials and Managers	5
1140	[1140] Senior Officials of Special-Interest Organisations	1
1141	[1141] Senior Officials of Political Party Organisations	0
1142	[1142] Senior Officials of Employers, Workers and Other Economic-Interest Organisations	11
1143	[1143] Senior Officials of Humanitarian and Other Special-Interest Organisations	0
1200	[1200] Corporate Managers	43
1210	[1210] Directors and Chief Executives	92
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	73
1223	[1223] Production and Operations Managers in Construction	0
...	(468 rows omitted)	13921
9212	[9212] Forestry Labourers	0
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	0
9312	[9312] Construction and Maintenance Labourers: Roads, Dams and Similar Constructions	3
9313	[9313] Building Construction Laborer	53
9320	[9320] Manufacturing Laborer	215
9330	[9330] Transport Lab., Freight Handler	154
-1	[-1] No Answer	70
-2	[-2] Does not apply	15256
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

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)). [This information can be related to a specific variable and is not necessarily generic.]
Hartmann/Schütz (2002): Die Klassifikation der Berufe und der Wirtschaftszweige im Sozio-oekonomischen Panel. Neuvercodung der Daten 1984–2001. Infratest Sozialforschung, München. https://www.diw.de/documents/dokumentenarchiv/17/diw_01.c.40132.de/vercodung.pdf
For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

isco08_$$ – Current Occupational Classification (ISCO-08) [generic]

110	[110] Commissioned Armed Forces Officers	6
210	[210] Non-Commissioned Armed Forces Officers	4
310	[310] Armed Forces Occupations, Other Ranks	27
1111	[1111] Legislators	6
1112	[1112] Senior Government Official	20
1113	[1113] Traditional Chiefs and Heads of Village	0
1114	[1114] Senior Officials of Special-Interest Organisations	10
1120	[1120] Managing Directors and Chief Executives	78
1211	[1211] Finance Managers	20
1212	[1212] Human Resource Managers	31
1213	[1213] Policy and Planning Managers	13
1219	[1219] Business Services and Administration Managers Not Elsewhere Classified	23
1221	[1221] Sales and Marketing Managers	62
1222	[1222] Advertising and Public Relations Managers	9
1223	[1223] Research and Development Managers	10
...	(412 rows omitted)	14135
9520	[9520] Street Vendors (excluding Food)	0
9611	[9611] Garbage and Recycling Collectors	14
9612	[9612] Refuse Sorters	8
9613	[9613] Sweeper, Related Laborer	3
9621	[9621] Messengers, Package Deliverers and Luggage Porters	60
9622	[9622] Odd Job Persons	0
9623	[9623] Meter Readers and Vending-Machine Collectors	0
9624	[9624] Water and Firewood Collectors	0
9629	[9629] Elementary Workers Not Elsewhere Classified	19
-1	[-1] No Answer	81

-2	[-2] Does not apply	15296
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Current occupation coded as ISCO-08. Tschersich and Schütz (2014, SOEP Survey Paper 470) provide detailed information on occupational coding. (Available in this classification since 2013 – wave bd.) [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

isei88_$$ – Last Reached Isei Value (International Socio-Economic Index) [generic]

16		916
19		113
20		314
21		90
22		17
23		504
24		54
25		820
26		170
27		103
28		72
29		793
30		1198
31		97
32		223
...	(33 rows omitted)	11585
69		1479
70		123
71		363
74		75
77		183
78		28
79		23
82		48
83		10
85		108
87		31
88		193
90		12
-1		38
-2		10152

Waves: all
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 occu-

pation. 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 Hendrickx which itself is based on Harry Gaanzeboom's SPSS algorithms.
Also available: occupational prestige scores (SIOPS, MPS) and occupational class (EGP). [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

kldb92_$$ – Current Occupational Classification (KldB92) [generic]

110	[110] Farmers, general	49
111	[111] Fruit and vegetable farmers (non-horticultural)	0
112	[112] Arable farmers (special, permanent crops)	0
113	[113] Livestock farmers and pasture farmers	0
114	[114] Seed, crop producers, propagators (non-horticultural)	1
115	[115] Crop protectors	1
116	[116] Farmers and landlords	0
118	[118] Farmers and wine growers	0
120	[120] Wine growers, general	2
121	[121] Vine propagators	0
129	[129] Other wine growers	0
130	[130] Agricultural workers, general	2
131	[131] Agricultural supervisors	0
132	[132] Agricultural machinery drivers	2
133	[133] Vineyard workers	1
...	(2261 rows omitted)	14532
9831	[9831] School leavers (job-searching) with (as yet) undefined occupation	1
9832	[9832] Other employees (job-searching) with (as yet) undefined occupation	1
9911	[9911] Specialized professionals without further specification	4
9921	[9921] Homeworkers without further specification	0
9931	[9931] Forepersons, group leaders without further specification	3
9941	[9941] Persons doing community service without further specification	0
9951	[9951] Self-employed persons without further specification	8
9961	[9961] Consultancy, planning professionals without further specification	6
9971	[9971] Other employees without further specification	70
-1	[-1] No Answer	10
-2	[-2] Does not apply	15242
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Current occupation coded as KldB92. Hartmann and Schütz (2002) provide detailed information on occupational coding. [This information can be related to a specific variable and is not necessarily generic.]
Hartmann/Schütz (2002): Die Klassifikation der Berufe und der Wirtschaftszweige im Sozio-oekonomischen Panel. Neuvercodung der Daten 1984–2001. Infratest Sozialforschung, München. https://www.diw.de/documents/dokumentenarchiv/17/diw_01.c.40132.de/vercodung.pdf
For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

kldb2010_$$ – Current Occupational Classification (KldB2010) [generic]

1104	[1104] Officer	5
1203	[1203] Senior Non-Commissioned Officers and Higher	5
1302	[1302] Junior Non-Commissioned Officers	0
1402	[1402] Armed Forces Personnel in Other Ranks	27
11101	[11101] Occupations in Farming (without Specialisation)-Unskilled/Semiskilled Tasks	12
11102	[11102] Occupations in Farming (without Specialisation)-Skilled Tasks	43
11103	[11103] Occupations in Farming (without Specialisation)-Complex Tasks	3
11104	[11104] Occupations in Farming (without Specialisation)-Highly Complex Tasks	11
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	1
11132	[11132] Technical Laboratory Occup. in Agriculture-Skilled Tasks	0
11133	[11133] Technical Laboratory Occup. in Agriculture-Complex Tasks	0
11182	[11182] Occupations in Farming (with Specialisation, Not Elsewhere Classified)-Skilled Tasks	0
...	(1262 rows omitted)	14410
94622	[94622] Prop Designers-Skilled Tasks	1
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	4
94712	[94712] Technical Occup. in Museums and Exhibitions-Skilled Tasks	0
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	101
-2	[-2] Does not apply	15312
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Current occupation coded as KldB2010. Tschersich and Schütz (2014, SOEP Survey Paper 471) provide detailed information on occupational coding. (Available in this classification since 2013 – wave bd.) [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

mps92_$$ – Last Reached Mps Value (Magnitude-Prestige-Skale, Wegener) [generic]

30	2
30.1000003814697	29
30.2000007629395	84
30.2999992370605	8
31	141
31.1000003814697	49
31.2000007629395	15

31.5		290
31.7000007629395		139
31.7999992370605		17
31.8999996185303		1
32		23
32.0999984741211		187
32.2000007629395		8
32.2999992370605		756
...	(155 rows omitted)	17091
123.900001525879		58
125.199996948242		12
132.100006103516		158
135.699996948242		97
138.199996948242		18
138.899993896484		10
139.800003051758		25
145.699996948242		105
152.5		150
153.5		6
191.300003051758		193
207.199996948242		38
216		33
-1		38
-2		10154

Waves: all
This variable gives the occupational prestige score developed by Wegener (1988) for all employed persons. Like the SIOPS prestige sore, Wegener's prestige scala measures a person's occupational prestige and was developed especially for use in the Federal Republic of Germany. MPS is assigned based on the German Federal Statistical Office's occupational classification of 1992 (KLDB92$$). The procedure has been documented in Frietsch and Wirth (2001).
Also available: occupational prestige scores (SIOPS, ISEI) and occupational class (EGP). [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

siops88_$$ – Last Reached Siops Value (Std. Internat. Occupational Prestige Scale) [generic]

13	18
15	90
17	13
18	1
19	352
20	359
21	930
22	233
23	272
24	50
25	292

26		42
27		18
28		151
29		173
...	(33 rows omitted)	15411
64		46
65		121
66		212
67		78
68		2
69		40
70		275
71		48
72		71
73		75
75		9
76		12
78		351
-1		38
-2		10152

Waves: all
This variable gives the occupational prestige score index for all employed persons. SIOPS$$ is based on ISCO-88 and was developed by Donald Treiman et al. The scale ranges from 6 to 78. The alogorithm is based on Fritsche and Wirth (2001).
Please also see occupational prestige scores (MPS$$), occupational status (ISEI$$), and occupational class (EGP$$). [This information can be related to a specific variable and is not necessarily generic.]
Frietsch, Rainer/Wirth, Heike (2001): Die Uebertragung der Magnitude-Prestigeskala von Wegener auf die Klassifikation der Berufe. In: ZUMA Nachrichten 48 (Jg.25): 139–165
For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

egp88_$$ – Last Reached Egp Value (Erikson, Goldthorpe, Portocarero) [generic]

1	[1] [I] Higher Managerial and Professional Workers	2664
2	[2] [II] Lower Managerial and Professional Workers	4802
3	[3] [IIIa] Routine Clerical Work	2552
4	[4] [IIIb] Routine Service and Sales Work	2770
5	[5] [IVa] Small Self-Employed With Employees	296
6	[6] [IVb] Small Self-Employed Without Employees	469
7	[7] [V] Manual Supervisors	0
8	[8] [VI] Skilled Manual Workers	2678
9	[9] [VIIa] Semi- and Unskilled Manual Workers	3185
10	[10] [VIIb] Agricultural Labour	262
11	[11] [IVc] Self-Employed Farmers	67
-1	[-1] No Answer	38
-2	[-2] Does not apply	10152
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
This variable gives the occupational class for all employed persons. EGP$$ is derived from the Standard International Socio-Economic Index of Occupational Status (ISEI). Technically, the variable was created by scaling the ISCO-88 classification. In addition, it is based on information about income, education and occupation. The EGP Index was documented by Ganzeboom/Treiman in 1996 and revised in 2003.
Information about supervisory status is only available from wave X (2007) on, the potential category (7) "Manual workers with supervisory status" is only assigned since then.
Annual information on the occupational position is used to generate the EGP-categories for the self-employed. In case no information on the number of employees is available, the EGP$$-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 EGP$$ 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.

Please also see occupational status (ISEI$$) and occupational prestige scores (SIOPS$$, MPS$$). [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

erljob$$ – Working In Occupation Trained For [generic]

1	[1] Yes	7978
2	[2] No	5219
3	[3] In Training	931
4	[4] Has No Job Training	821
-1	[-1] No Answer	1218
-2	[-2] Does not apply	13768
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
This variable is designed to offer annual data on all employed persons, indicating whether they are working in the occupation they were trained for. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

ausb$$ – Required Training For Job [generic]

1	[1] No Training	1436
2	[2] Intro. To Job	1727
3	[3] On-The-Job Training	551
4	[4] Courses	312
5	[5] Vocational Training	7026

6	[6] Technical School, Engineering (East) 90-96	0
7	[7] Technical College, University until 1998	0
8	[8] Technical College since 1999	1392
9	[9] University since 1999	2184
-1	[-1] No Answer	1539
-2	[-2] Does not apply	13768
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
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 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.
The answer option "completed technical school" was only used from 1990 to 1993 in the East German version of the questionnaire. Since not all employed people are asked the question about required training every year, the value (6) of the variable AUSB$$ is valid up to 1996.
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. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

$erwzeit – Length Of Time With Firm [generic]

Waves: all
The variable $ERWZEIT 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, if it can be done under reasonable assumptions.

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). [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

$tatzeit – Actual Work Time Per Week [generic]

Waves: all
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 $VEBZEIT and $UEBSTD. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

$vebzeit – Agreed Upon Work Time Per Week [generic]

Waves: all
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.
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 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 $TATZEIT and $UEBSTD. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

$uebstd – Overtime Per Week [generic]

Waves: all
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 are 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 $UEBSTD 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 years 1984, 1985 and 1987, respondents were not asked about number of hours of overtime per week. The variable $UEBSTD for these years 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 off and gives the number of overtime hours as a decimal number.
Please also see $VEBZEIT and $TATZEIT. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

oeffd$$ – Civil Service [generic]

1	[1] Yes	3534
2	[2] No	11002
-1	[-1] No Answer	1631
-2	[-2] Does not apply	13768
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
Status variable: This variable is designed to provide annual data on employment in the civil service for all employed persons. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

nace$$ – Industry Occupation [pbra] (NACE Rev. 1.1, Sector) [generic]

1	[1] Agriculture, Hunting, Related Service Activities	220
2	[2] Forestry, Logging, Related Service activities	21
5	[5] Fishing, Operation Of Fish Hatcheries And Fish Farms	1
10	[10] Mining Of Coal And Lignite; Extraction Of Peat	8
11	[11] Extraction Of Crude Petroleum And Natural Gas	7

12	[12] Mining Of Uranium And Thorium Ores	0
13	[13] Mining Of Metal Ores	0
14	[14] Other Mining And Quarrying	2
15	[15] Manuf Food Products And Beverages	322
16	[16] Manuf Tobacco Products	2
17	[17] Manuf Textiles	46
18	[18] Manuf Wearing Apparel; Dressing And Dyeing Of Fur	30
19	[19] Tanning,Dressing Of Leather; Manuf luggage, Footwear	7
20	[20] Manuf Wood Products, Except Furniture	54
21	[21] Manuf Pulp, Paper And Paper Products	49
...	(38 rows omitted)	10682
85	[85] Health And Social Work	2237
90	[90] Sewage And Refuse Disposal, Sanitation And Related	64
91	[91] Activities Of Membership Organizations NEC.	169
92	[92] Recreational, Cultural And Sporting Activities	303
93	[93] Other Service Activities	167
95	[95] Private Households With Employed Persons	70
96	[96] Industry - NEC	0
97	[97] Handcraft, Trade - NEC	0
99	[99] Extra-territorial Organizations And Bodies	8
-1	[-1] No Answer	70
-2	[-2] Does not apply	15396
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
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 Rev. 1.1 which is on this level similar to Rev. 1). Respondents answer the question in their own words regarding the industry in which they are currently working. In order to facilitate international comparability, the European industry standard classification system is used by Kantar Public using Level 2 (Divisions) of the classification system. The classification 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). [This information can be related to a specific variable and is not necessarily generic.]
Detailed description: Hartmann/Schütz (2002): Die Klassifikation der Berufe und der Wirtschaftszweige im Sozio-oekonomischen Panel. Neuvercodung der Daten 1984–2001. Infratest Sozialforschung, München. https://www.diw.de/documents/dokumentenarchiv/17/diw_01.c.40132.de/vercodung.pdf
For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

nace2$$ – Industry Occupation [pbra] (NACE Rev. 2, Sector) [generic]

1	[1] Crop and animal production, hunting and related service activities	155
2	[2] Forestry and logging	21
3	[3] Fishing and aquaculture	1
5	[5] Mining of coal and lignite	7

6	[6] Extraction Of Crude Petroleum And Natural Gas	7
7	[7] Mining Of Metal Ores	0
8	[8] Other Mining And Quarrying	4
9	[9] Mining support service activities	0
10	[10] Manufacture of food products	293
11	[11] Manufacture of beverages	27
12	[12] Manuf Tobacco Products	2
13	[13] Manuf Textiles	46
14	[14] Manuf Wearing Apparel; Dressing And Dyeing Of Fur	30
15	[15] Manufacture of leather and related products	7
16	[16] Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	54
...	(64 rows omitted)	13243
91	[91] Libraries, archives, museums and other cultural activities	23
92	[92] Gambling and betting activities	12
93	[93] Sports activities and amusement and recreation activities	74
94	[94] Activities of membership organisations	169
95	[95] Repair of computers and personal and household goods	11
96	[96] Other personal service activities	156
97	[97] Private Households With Employed Persons	71
98	[98] Herstellung von Waren und Erbringung von Dienstleistungen durch private Hausha	0
99	[99] Extra-territorial Organizations And Bodies	8
-1	[-1] No Answer	70
-2	[-2] Does not apply	15444
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
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 Rev. 2). Respondents answer the question in their own words regarding the industry in which they are currently working. In order to facilitate international comparability, the European industry standard classification system is used by Kantar Public using Level 3 (Divisions) of the classification system. Tschersich and Schütz (2014, SOEP Survey Paper 471) provide detailed information on this coding. (Available in this classification since 2013 – wave bd.) [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

betr$$ – Size of the Company [generic]

1	[1] Lt 5	1429
2	[2] Ge 5 Lt 10	1360
3	[3] Ge 11 Lt 20	1277
4	[4] Until 90: Lt 20	0
5	[5] 91-04: Ge 5 Lt 20	0
6	[6] Ge 20 Lt 100	2558

7	[7] Ge 100 Lt 200	1281
8	[8] Until 98: Ge 20 Lt 200	0
9	[9] Ge 200 Lt 2000	2959
10	[10] Ge 2000	3846
11	[11] Self-Employed Without Coworkers	481
-1	[-1] No Answer	976
-2	[-2] Does not apply	13768
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
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. In the years 1991, 1999, and 2005, respondents were asked about company size in a more detailed form, so when the data were recalculated for the years 1984–90, the original codes (1)–(5) were changed to (4)(8)(9)(10) and (11). For the years 1999 to 2004, Code (8) was differentiated into (6) and (7). From 2005 on, the prior category "5 to 20 employees" (5) has been split into the two categories "5 to 10 employees" (2) and "11 to 20 employees" (3).
In 2012, the questionnaire provides a one-time-only information on the size of the local establishment in addition to the size of the entire company (BETR$$). The enriched questionnaire revealed that in previous interviews, persons have mistakenly provided information on the local establishment size instead of the entire company size, especially if their entire company had 2000 and more employees. Due to the importance of longitudinal consistency, these persons where identified, and their 2012 original value of the entire company size was replaced by their value of the local establishment size. Furthermore, data of persons without a job change were modified if:

- the entire company size varies from 2011 to 2012,
- the local establishment size in 2012 matches the entire company size in 2011.

Those person's values of the entire company in 2012 were replaced by their values of the local establishment size in 2012.
Please also see ALLBET$$ for a broader categorization of the firm size, which is appropriate for analyses that include all sample years. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

allbet$$ – Core Category Size Of The Company [generic]

1	[1] LT 20	4066
2	[2] GE 20 LT 200	3839
3	[3] GE 200 LT 2000	2959
4	[4] GE 2000	3846
5	[5] Self-Employed Without Coworkers	481
-1	[-1] No Answer	976
-2	[-2] Does not apply	13768
-3	[-3] Answer improbable	0

-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
This variable is designed to provide annual data on the core size category of the company for all employed persons. Since respondents were asked about company size in more detailed form in the years 1991, 1999, and 2005 (see also BETR$$), the variable ALLBET$$ contains the lowest common denominator of the variable BETR$$, i.e., the firm size categories available across all SOEP waves. This broader categorization corresponds to the values of variables BETR84 to BETR90 and offers a variable that is consistent across all waves.
In order to maintain longitudinal consistency, there were modifications after a one-time change in the questionnaire in 2012. Please also see BETR$$ for more information. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

8 Last Occupation

jobend$$ – Reasons for occupational change [generic]

1	[1] Terminated by employer	304
2	[2] Temporary contract expired	0
3	[3] Education, training completed	0
4	[4] Own resignation	653
5	[5] Mutual termination	158
6	[6] Employee requested transfer	0
7	[7] Company transferred employee	0
8	[8] Ended self-employment	57
9	[9] Temporary contract expired or education/training completed	286
10	[10] Took early retirement	0
11	[11] Company closed down	98
12	[12] Old-age pension	145
13	[13] Leave of absence/sabbatical (1999-2010)	0
14	[14] Leave, maternity leave and parental leave (1991-1998), since 2011	211
15	[15] Other incl. early retirement, company closed, old-age pension, leave of absence/sabbatical (1985-1986)	0
16	[16] Other incl. company closed, old-age pension,leave of absence/sabbatical (1987-1990)	0
17	[17] Other incl. mutual termination (1991-1998)	0
-1	[-1] No Answer	68
-2	[-2] Does not apply	27420
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	535
-6	[-6] Version of questionnaire with modified filtering	0

Waves: -1985
This variable is designed to offer annual data on reasons for an occupational change for all formerly employed persons, persons with a job change or persons on leave. For years 1985–1990 also persons who changed positions in the same company are considered. Only persons with valid dates for an occupational or positional change are included. Likewise to the ques-

tionnaire the variable offers data from interview date to interview date not from one year to the following. Respondents are asked about their annual and possibly same occupational change in two consecutive interviews, duplicate answers are therefore considered only once and the older statement is dominant. If a respondent stated a job termination in the current interview which was before the interview date in the previous year but didn't reported this in the previous interview this termination has been counted for the current interview. For years 1985 up to 1998 every given reason was coded as separate variable with variable values "Yes" (1) and "Does not apply" (-2), which resulted in up to 13 different variables. Since 1999 all given reasons have been collected in one single variable with diverse values. If respondent states more than one reason for job termination, the first of the stated reasons is dominant. The year 1990 is specific due to introduction of sample for East Germany. Since the questionnaire did not contain the information on reasons for end of the job all the observations in Sample C in year 1990 obtained value (-5).
Please pay attention to special codes (15), (16), and (17)! These codes were necessary due to the variety of the given values over the years. In any years respondents were asked about reasons for change with more or less given answers and from years 1985–1998 also the answer "Other" was possible. While all explicit reasons have been recoded to uniform values, the answer "Other" then in some years includes reasons for which in other years was separately asked for: "Other" was coded (15) for years 1985 and 1986, (16) for years 1987–1990 and (17) for years 1991–1998.
For years 1991–1998 and 2011–2012 there is a variable value "Leave, maternity leave and parental leave" whereas for 1999–2010 the given reason covered only "Leave of absence/sabbatical". Note that codes (2) and (3) for years 1985–1998 have been merged to code (9) since 1999.
In 2013 the respondents of the newly introduced migration sample were not asked about an occupational change, but in 2014 they were. So information on jobend for the migration sample in 2013 were taken from the questionnaire in 2014.
Since 2006 youth questionnaires have been embodied in survey tools. Nevertheless, the questionnaires do not contain the information on jobend. Therefore, in order to account for the persons represented by the youth questionnaires the negative value (-5) has been introduced starting from year 2006. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

9 Employment History

expft$$ – Working Experience Full-Time Employment [generic]

Waves: all
This variable reflects the total length of full-time employment in the respondent's career up to the point of the interview in a given year. 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). EXPFT$$ 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 without annual data (not contained in PBIOSPE) are only

assigned a non-missing value for this variable if they joined SOEP by the age of 18 and if there is calendar data on them in ARTKALEN.
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).
Please also see EXPPT$$ and EXPUE$$. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

exppt$$ – Working Experience Part-Time Employment [generic]

0		14769
0.100000001490116		137
0.200000002980232		151
0.300000011920929		319
0.400000005960464		126
0.5		670
0.600000023841858		174
0.699999988079071		128
0.800000011920929		235
0.899999976158142		104
1		993
1.10000002384186		112
1.20000004768372		114
1.29999995231628		200
1.39999997615814		108
...	(358 rows omitted)	10491
43.7999992370605		1
44		1
44.5		1
45.2999992370605		1
45.4000015258789		1
45.5		1
45.7000007629395		1
45.7999992370605		1
46.2000007629395		1
47		2
47.2999992370605		1
48.0999984741211		1
49.7000007629395		1
50		1
-1		1089

Waves: all
This variable reflects the total length of part-time employment in the respondent's career up to the point of the interview in a given year. 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). EXPPT$$ 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 without annual data (not contained in PBIOSPE) are only assigned a non-missing value for this variable if they joined SOEP by the age of 18 and if there is calendar data on them in ARTKALEN.
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).
Please also see EXPFT$$ and EXPUE$$. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

expue$$ – Unemployment Experience [generic]

0		18346
0.100000001490116		343
0.200000002980232		315
0.300000011920929		398
0.400000005960464		176
0.5		1686
0.600000023841858		194
0.699999988079071		184
0.800000011920929		266
0.899999976158142		132
1		1162
1.10000002384186		148
1.20000004768372		124
1.29999995231628		169
1.39999997615814		100
...	(211 rows omitted)	5082
25.8999996185303		2
26		1
26.1000003814697		1
26.2999992370605		1
26.8999996185303		1
27		3
27.2000007629395		1
28		2
29		3
29.2999992370605		1
32		2
34		1
36		1
40		1
-1		1089

Waves: all

This variable reflects the total length of unemployment in the respondent's career up to the point of the interview in a given year. 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). EXPUE$$ gives the length of time in years with months in decimal form.
If there is no monthly calendar data available on a given year in a respondent's career, the annual data from PBIOSPE is used for that year. In the most current wave the variable only uses up-to-date information from the newly answered Biography Questionnaires. If the year in which a spell started and ended is the same, and if there is no monthly data, a spell of 0.5 years is assumed. Persons without annual data (not contained in PBIOSPE) are only assigned a non-missing value for this variable if they joined SOEP by the age of 18 and if there is calendar data on them in ARTKALEN.
Persons whose life course has been observed completely but with no spell of full-time employment are assigned the code (0). The code (-1) is assigned to all persons whose life course has not been observed completely. Persons with inconsistent information receive a (-3).
Please also see EXPFT$$ and EXPPT$$. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

10 School, Higher and Vocational Education

isced97_$$ – ISCED-1997-Classification [generic]

0	[0] in school	378
1	[1] inadequately	2135
2	[2] general elementary	3712
3	[3] middle vocational	11826
4	[4] vocational + Abi	2134
5	[5] higher vocational	1414
6	[6] higher education	6698
-1	[-1] No Answer	1638
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
The educational variable ($ISCED97) classifies all correspondents' educational degrees according to the "International Standard Classification of Education (ISCED)" of 1997 in order to make degrees internationally comparable. The variable is generated retrospectively from 1984 onwards taking into account degrees and diplomas attained in both general schooling and in vocational/university education and indicates the highest degree obtained. E.g., persons who did not indicate secondary school degrees/diplomas but 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 earlier waves – in particular on PhD – we include all tertiary degrees in 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. OECD (1999): Classifying Educational Programmes: Manual for ISCED-97 Implementation in OECD Countries, Paris. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

isced11_$$ – ISCED-2011-Classification [generic]

0	[0] in school	378
1	[1] Primary education	2180
2	[2] Lower secondary education	3877
3	[3] Upper secondary education	11890
4	[4] Post-secondary non-tertiary education	2122
5	[5] Short-cycle tertiary education	1044
6	[6] Bachelors or equivalent level	4527
7	[7] Masters or equivalent level	1973
8	[8] Doctoral or equivalent level	293
-1	[-1] No Answer	1651
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: 2010–
The educational variable ($ISCED11) classifies all correspondents' educational degrees according to the "International Standard Classification of Education (ISCED)" of 2011 in order to make degrees internationally comparable. The variable is generated retrospectively From 2010 onwards taking into account degrees and diplomas attained in both general schooling and in vocational/university education and indicates the highest degree obtained.
[This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

casmin$$ – CASMIN Classification [generic]

0	[0] (0) In School	379
1	[1] (1a) Inadequately Completed	2374
2	[2] (1b) General Elementary School	2788
3	[3] (1c) Basic Vocational Qualification	5125
4	[4] (2b) Intermediate General Qualification	847
5	[5] (2a) Intermediate Vocational	5325
6	[6] (2c_gen) General Maturity Certificate	0
7	[7] (2c_voc) Vocational Maturity Certificate	4553
8	[8] (3a) Lower Tertiary Education	1838
9	[9] (3b) Higher Tertiary Education	4811
-1	[-1] No Answer	1895
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
Another internationally comparable educational variable is $CASMIN where educational degrees/diplomas are classified according to the scheme "Comparative Analysis of Social Mobility in Industrial Nations (CASMIN)". As for $ISCED, the variable is generated for all respondents retroactively from 1984 onwards and indicates the highest degree obtained by the respondent. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

$bilzeit – Amount Of Education Or Training In Years [generic]

7	1569
8.5	56
9	2969
10	2132
10.5	3549
11	1326
11.5	3848
12	2491
13	1302
13.5	379
14	522
14.5	647
15	2523
16	803
17	97
18	2742
-1	2523
-2	457

Waves: all
The following statements describe the standard computation for schooling (including years of secondary vocational education). As can be seen, the code is not very differentiated. For example, special schools for health care professions and other kinds of specialized schools are all included in the "technical school" label. However, in Germany, this code is the most commonly used one when earnings functions based on human capital theory are estimated. $BILZEIT is now computed for all samples. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

$psbil – School-Leaving Degree [generic]

1	[1] Secondary School Degree	5890
2	[2] Intermediate School Degree	6750
3	[3] Technical School Degree	1425
4	[4] Upper Secondary Degree	5432
5	[5] Other Degree	5758
6	[6] Dropout, No School Degree	1700
7	[7] No School Degree Yet	458

8	[8] No School Attended	491
-1	[-1] No Answer	2031
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
All respondents in all SOEP subsamples are asked about diplomas/degrees attained for completion of secondary/tertiary education (1984–1993 blue questionnaire; since 1994 biographical questionnaire) the first time they participate in SOEP. First: to generate this variable, the different diploma/degree categories provided for Subsamples B and D (see $PSBILA) as well as C (see $PSBILO) are integrated into the West German diploma/degree categories (Subsample A) and continued on in this form. Second: this data is regularly updated to take into account any changes in highest diploma/degree attained. With the survey of 2000, all educational information was collected again and is reflected in the variables. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

$pbbil01 – Vocational Degree Received [generic]

1	[1] Apprenticeship	9170
2	[2] Vocational School	2109
3	[3] Health Care School	180
4	[4] Technical School	1370
5	[5] Civil Service Training	500
6	[6] Other degree	1441
7	[7] Completed Vocational Training/Education in Germany	229
-1	[-1] No Answer	669
-2	[-2] Does not apply	14267
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
All respondents in all subsamples are asked about vocational degrees attained the first time they participate in SOEP (1984–1993 blue questionnaire; since 1994 biographical questionnaire). To generate the variable, the different vocational degrees for Subsamples B and D (cf. $PBBILA) as well as C (cf. $PBBILO) are integrated into the West German vocational degree categories (Subsample A). 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. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

$pbbil02 – College Degree [generic]

1	[1] Technical College	1633

2	[2] University, Technical College	2841
3	[3] College Not In Germany	1499
4	[4] Engineering, Techncial School (East)	206
5	[5] University (East)	174
6	[6] graduation, state doctorate	199
7	[7] graduation, state doctorate (foreign country, east)	94
8	[8] institution of higher education (youth)	0
9	[9] Dual Studies, University of Cooperative Education	36
10	[10] Other Colleges (since 2014)	16
-1	[-1] No Answer	669
-2	[-2] Does not apply	22568
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
All respondents in all subsamples are asked about completed college education the first time they participate in SOEP (1984–1993 blue questionnaire; since 1994 biographical questionnaire). To generate the variable, the different degrees/dimplomas for all subsamples are integrated. Category (3) "college abroad" is only defined for persons who completed a foreign-language version of the questionnaire (mainly persons from Samples B and D). Generation of the variable entails combining the categories to make them compatible with the annual question about changes in vocational degrees/diplomas attained. Since 2002, there have been two separate codes (4 and 5) for degrees/diplomas attained in the former GDR. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

$pbbil03 – No Vocational Degree [generic]

1	[1] No Vocation Degree	8081
2	[2] Apprenticeship	984
3	[3] University	954
-1	[-1] No Answer	669
-2	[-2] Does not apply	19247
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
In connection with the question about vocational degrees ($PBBIL01 and $PBBIL02), 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. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

$psbilo – School-Leaving Degree East Germany [generic]

1	[1] 8th Grade Completed	890
2	[2] 10th Grade Completed	1990
3	[3] College Entrance Exam	739
4	[4] Other Degree	31
5	[5] Dropout, No School Degree	22
6	[6] No School Degree Yet	0
-1	[-1] No Answer	0
-2	[-2] Does not apply	26263
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
As a supplement to the variable $PSBIL the highest secondary school degree/diploma in East Germany is provided as a separate variable and updated if necessary for 1991. Since 1992, secondary degrees/diplomas are asked only in the West German version. 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. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

$pbbilo – Vocational Degree Received East Germany [generic]

1	[1] Vocational Training	1505
2	[2] Master Craftsman	152
3	[3] Engineering, Technical Degree	330
4	[4] Other degree	24
-1	[-1] No Answer	0
-2	[-2] Does not apply	27924
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
To supplement the variable $PBBIL01 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 $PBBIL03 if this degree has not been replaced by a more recently attained, higher-level university or college degree). [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

$psbila – School-Leaving Degree Outside Germany [generic]

1	[1] School, No Degree	1477
2	[2] School, With Degree	2282
3	[3] Vocational Extension School	3284
4	[4] School Leaving Degree[Sbil] Acquired Abroad	1
5	[5] Certificate from a different school	117
-1	[-1] No Answer	0
-2	[-2] Does not apply	22774
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
As a supplement to the $PSBIL, this variable provides annually updated data on the highest secondary school degree/diploma attained abroad. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

$pbbila – Vocational Degree Outside Germany [generic]

1	[1] On-The-Job Training	76
2	[2] Vocational Training	117
3	[3] Vocational School	158
4	[4] College	557
5	[5] Other	33
6	[6] Vocational Degree[Bbil01] Acquired Abroad	0
7	[7] College Education[Bbil02] Acquired Abroad	0
8	[8] Completed Vocational Training/Education Other Country	357
9	[9] graduation, state doctorate (foreign country)	30
11	[11] With Certificate, On-The-Job Training	92
12	[12] With Certificate, Vocational Training	379
13	[13] With Certificate,Vocational School	509
14	[14] With Certificate, University	1198
15	[15] With Certificate, Other	86
16	[16] With Certificate, Vocational Degree[bbil01] Acquired Abroad	6
17	[17] With Certificate, College Education[Bbil02] Acquired Abroad	16
18	[18] With Certificate, Completed Vocational Training/Education Other Country	0
19	[19] With Certificate, PhD (Other Country)	68
-1	[-1] No Answer	0
-2	[-2] Does not apply	26253
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
As a supplement to the variable $PBBIL01, this variable gives (and updates) the highest-level vocational degree attained abroad. [This information can be related to a specific variable and

is not necessarily generic.]
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

field$$ – Field of tertiary education [generic]

1	[1] Sprach- und Kulturwissenschaften allgemein	35
2	[2] Protestant Theology	51
3	[3] Catholic Theology	2
4	[4] Philosophy	18
5	[5] History	30
6	[6] Bibliothekswissenschaft, Dokumentation, Publizistik	14
7	[7] Allgemeine und vergleichende Literatur und Sprachwissenschaft	49
8	[8] Altphilologie (klass. Philologie), Neugriechisch	5
9	[9] Germanistik (Deutsch, germanische Sprachen ohne Anglistik)	99
10	[10] English Studies	40
11	[11] Romance Philology	20
12	[12] Slavonic Studies	9
13	[13] Aussereuropaeische Sprach- und Kulturwissenschaften	7
14	[14] Cultural Studies	8
15	[15] Psychology	71
...	(35 rows omitted)	3162
67	[67] Regional Planning	3
68	[68] Civil Engineering	118
69	[69] Surveying and Mapping	11
74	[74] Art, Aesthetics	23
75	[75] Fine Arts	7
76	[76] Design	39
77	[77] Darstellende Kunst, Film und Fernsehen, Theaterwissenschaft	11
78	[78] Music, Musicology	46
83	[83] Outside the structure of the university system	14
-1	[-1] No Answer	144
-2	[-2] Does not apply	25874
-3	[-3] Answer improbable	25
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: -1985
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 $PBBIL02. While the latter variable records if a person holds a degree FIELD$$ contains more detailed information on the type of the degree. The data of the generated variable FIELD$$ 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 FIELD$$.
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 FIELD$$ we combine these two types of information. However, since the retrospective information was not collected before 2001 the variable covers until 2000 only persons for whom we have prospectively observed the end of study. This explains why the number of valid observations is rather small in these years. Information on the data source is stored in the variable FDT_F$$.
Each year the variable contains the most recently collected information. Take for instance a person for whom we have observed a first degree in sociology in 1987 and a second degree in economics in 1991. For this person the variables FIELD$$ would be filled as follows:

year	value
1984–1986	-2 does not apply
1987–1990	26 political/social science
1991–today	30 economics

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 FIELD$$ we recoded the original values. As the newer version is more precise this could be done with hardly any loss of information. Some categories are collapsed. Category 3 is coded as 2 (no distinction between catholic and protestant theology), 14 as 13, 17 as 16, 24 as 23, 25 as 26 and 48 as 49. The original values of the data collected from the person questionnaire up to 2009 are stored in the respective variables in the dataset $P.
Please note that for respondents from the newly introduced migration sample a valid value to FIELD$$ was only assigned if the tertiary degree was attained in Germany. [This information can be related to a specific variable and is not necessarily generic.]
Statistisches Bundesamt (2009): Bildung und Kultur. Studierende an Hochschulen, Fachserie 11, Reihe 4.1, Wiesbaden: 446ff, Übersicht 1: “Fächergruppen, Studienbereiche und Studienfächer”.
For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

degree$$ – Type of tertiary degree [generic]

11	[11] Magister	125
12	[12] Diplom (University)	908
13	[13] Bachelor	164
14	[14] Master	142
15	[15] 1st State Examination	122

16	[16] Other state examination	126
21	[21] Diplom (at technical college, technical college for administration)	660
22	[22] Bachelor (at technical college, technical college for administration)	126
23	[23] Master (at technical college, technical college for administration)	31
31	[31] Teacher training,BA,MA at elementary, lower secondary schools/primary level	112
32	[32] Teacher training,BA,MA at 2ndary level 1/elementary schools/primary level	6
33	[33] Teacher training,BA,MA at intermediate scndry schools/scndry level I	48
34	[34] Teacher training,BA,MA at secondary level II and I	3
35	[35] Teacher training,BA,MA at academic 2ndry schools,2ndry levl 2,genrl school	60
36	[36] Teacher training,BA,MA at special needs schools	34
37	[37] Teacher training,BA,MA at vocational schools	12
38	[38] Teacher training, other	339
41	[41] Academic degree in the arts	15
42	[42] Doctorate	202
43	[43] Post-doctoral dissertation (Habilitation)	6
44	[44] Other degree	135
98	[98] Not categorizable	30
-1	[-1] No Answer	146
-2	[-2] Does not apply	26368
-3	[-3] Answer improbable	15
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: -1985

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 $PBBIL02. While the latter variable records if a persons holds a degree, DEGREE$$ contains more detailed information on the type of the degree. The data of the generated variable DEGREE$$ 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 DEGREE$$.

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 DEGREE$$ we combine these two types of information. However, since the retrospective information was not collected before 2001 the variable covers until 2000 only persons for whom we have prospectively observed the end of study. This explains why the number of valid observations is rather small in these years. Information on the data source is stored in the variable FDT_F$$.

Each year the variable contains the most recently collected information. Take for instance a person for whom we have observed first an applied university diploma in 1987 and a university diploma in 1991. For this person the variables DEGREE$$ would be filled as follows:

year	value
1984–1986	-2 does not apply
1987–1990	21 diploma (applied university)
1991–today	12 diploma (university)

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). Until 2009, data from the person questionnaire were coded using an earlier version of this classification (1982) which was slightly revised in 2009 (inclusion of Bachelor and Master degrees). Since 2010 the data were coded according to the classification presented here. In the variable DEGREE$$ we recoded the original values from years 2009 and earlier. As the newer version is more precise this could be done with hardly any loss of information. Some categories are collapsed. Category 16 was mostly likely coded as 15 in earlier years, 34 as 35 and 43 as 44. The original values of the data collected from the person questionnaire up to 2009 are stored in the respective variables in the dataset $P.
Please note that for respondents from the newly introduced migration sample a valid value to DEGREE$$ was only assigned if the tertiary degree was attained in Germany. [This information can be related to a specific variable and is not necessarily generic.]
Statistiches Bundesamt (2009): Bildung und Kultur. Studierende an Hochschulen, Fachserie 11, Reihe 4.1, Wiesbaden: 449ff, Übersicht 2: „Prüfungsgruppen und Abschlussprüfungen".
For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

traina$$ – Apprenticeship - two-digit occupation KldB92 [generic]

110	[110] Farmers, general	78
111	[111] Fruit and vegetable farmers (non-horticultural)	0
112	[112] Arable farmers (special, permanent crops)	2
113	[113] Livestock farmers and pasture farmers	0
114	[114] Seed, crop producers, propagators (non-horticultural)	1
115	[115] Crop protectors	2
116	[116] Farmers and landlords	3
118	[118] Farmers and wine growers	0
120	[120] Wine growers, general	4
121	[121] Vine propagators	0
129	[129] Other wine growers	0
130	[130] Agricultural workers, general	17
131	[131] Agricultural supervisors	0
132	[132] Agricultural machinery drivers	0
133	[133] Vineyard workers	0

...	(2261 rows omitted)	8620
9831	[9831] School leavers (job-searching) with (as yet) undefined occupation	0
9832	[9832] Other employees (job-searching) with (as yet) undefined occupation	0
9911	[9911] Specialized professionals without further specification	4
9921	[9921] Homeworkers without further specification	0
9931	[9931] Forepersons, group leaders without further specification	0
9941	[9941] Persons doing community service without further specification	0
9951	[9951] Self-employed persons without further specification	0
9961	[9961] Consultancy, planning professionals without further specification	2
9971	[9971] Other employees without further specification	255
-1	[-1] No Answer	107
-2	[-2] Does not apply	20840
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: -1985

The variable is designed to provide information on the occupation of vocational training which adds details to the information recorded in the variable $PBBIL01. In addition to the variable TRAINA$$, we provide the variables TRAINB$$, TRAINC$$ and TRAIND$$. All these variables record the occupation of vocational training. The difference is that TRAINA$$ contains information on vocational training within the German dual system which combines firm-based and school-based training (apprenticeship). TRAINB$$ is designed to provide information on the occupation of full-time school based vocational training. TRAINC$$ contains information on level vocational training (e.g., Meister, Techniker). TRAIND$$ is designed to provide information on the occupation of civil servant training ("Beamtenausbildung"). We describe in brief detail the construction of the variable TRAINA$$. TRAINB$$, TRAINC$$ and TRAIND$$ are constructed in an analogous manner.

The data of the generated variable TRAINA$$ 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 TRAINA$$.

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 vocational qualifications but to qualifications obtained during the time before being part of the SOEP sample.

In the variable TRAINA$$ we combine these two types of information. However, since the retrospective information was not collected before 2001 the variable covers until 2000 only persons for whom we have prospectively observed the end of study. This explains why the number of valid observations is rather small in these years. Information on the data source is stored in the variable FDT_F$$.

Each year the variable contains the most recently collected information. Take for instance a person for whom we have observed a first vocational qualification as an electrician in 1987 and a second qualification as a car mechanic in 1991. For this person the variables TRAINA$$ would be filled as follows:

year	value
1984–1986	-2 does not apply
1987–1990	31 electrical occupation
1991–today	28 automotive/flight industry occupation

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.
The variable is coded according to the classification of occupations at two-digit level („Berufsgruppen") provided by the Statistisches Bundesamt (1992). Other SOEP occupation variables are coded at four-digit level. The reason why the variable TRAINA$$ is provided at two-digit level only is that until 2009 the data from the two different sources were coded according two different classifications which could be combined at a higher level of aggregation only. The person questionnaire data were coded according to the classification of occupations provided by the Bundesanstalt für Arbeit (1988, four-digit level) while the biography data use the classification provided by the Statistisches Bundesamt (1992, four-digit level). Since 2010 both types of data are coded according to the latter classification. The four-digit version of the biography data can be provided upon request. The original values of the data collected from the person questionnaire up to 2009 are stored in the respective variables in the dataset $P.
Please note that for respondents from the newly introduced migration sample a valid value to TRAINA$$ was only assigned if the vocational training was completed in Germany. [This information can be related to a specific variable and is not necessarily generic.]
Hartmann/Schütz (2002): Die Klassifikation der Berufe und der Wirtschaftszweige im Sozio-oekonomischen Panel. Neuvercodung der Daten 1984–2001. Infratest Sozialforschung, München.
For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

trainb$$ – Vocational school - two-digit occupation KldB92 [generic]

110	[110] Farmers, general	1
111	[111] Fruit and vegetable farmers (non-horticultural)	0
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	0
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	1
131	[131] Agricultural supervisors	0
132	[132] Agricultural machinery drivers	0
133	[133] Vineyard workers	0
...	(2261 rows omitted)	1643
9831	[9831] School leavers (job-searching) with (as yet) undefined occupation	0

9832	[9832] Other employees (job-searching) with (as yet) undefined occupation	0
9911	[9911] Specialized professionals without further specification	2
9921	[9921] Homeworkers without further specification	0
9931	[9931] Forepersons, group leaders without further specification	0
9941	[9941] Persons doing community service without further specification	0
9951	[9951] Self-employed persons without further specification	0
9961	[9961] Consultancy, planning professionals without further specification	0
9971	[9971] Other employees without further specification	115
-1	[-1] No Answer	44
-2	[-2] Does not apply	28129
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: -1985
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 TRAINA$$ for more details on the construction and the values of the variable. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

trainc$$ – Higher vocational school - two-digit occupation KldB92 [generic]

110	[110] Farmers, general	22
111	[111] Fruit and vegetable farmers (non-horticultural)	0
112	[112] Arable farmers (special, permanent crops)	2
113	[113] Livestock farmers and pasture farmers	0
114	[114] Seed, crop producers, propagators (non-horticultural)	0
115	[115] Crop protectors	0
116	[116] Farmers and landlords	0
118	[118] Farmers and wine growers	0
120	[120] Wine growers, general	1
121	[121] Vine propagators	0
129	[129] Other wine growers	0
130	[130] Agricultural workers, general	0
131	[131] Agricultural supervisors	0
132	[132] Agricultural machinery drivers	0
133	[133] Vineyard workers	0
...	(2261 rows omitted)	1001
9831	[9831] School leavers (job-searching) with (as yet) undefined occupation	0
9832	[9832] Other employees (job-searching) with (as yet) undefined occupation	0
9911	[9911] Specialized professionals without further specification	0
9921	[9921] Homeworkers without further specification	0
9931	[9931] Forepersons, group leaders without further specification	0
9941	[9941] Persons doing community service without further specification	0
9951	[9951] Self-employed persons without further specification	0
9961	[9961] Consultancy, planning professionals without further specification	0
9971	[9971] Other employees without further specification	76

-1	[-1] No Answer	25
-2	[-2] Does not apply	28808
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: -1985
The variable is designed to provide information on the occupation of higher level vocational training (e.g., Meister, Techniker). See the description of variable TRAINA$$ for more details on the construction and the values of the variable. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

traind$$ – Civil servant training - two-digit occupation KldB92 [generic]

110	[110] Farmers, general	0
111	[111] Fruit and vegetable farmers (non-horticultural)	0
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	0
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	0
131	[131] Agricultural supervisors	0
132	[132] Agricultural machinery drivers	0
133	[133] Vineyard workers	0
...	(2261 rows omitted)	348
9831	[9831] School leavers (job-searching) with (as yet) undefined occupation	0
9832	[9832] Other employees (job-searching) with (as yet) undefined occupation	0
9911	[9911] Specialized professionals without further specification	0
9921	[9921] Homeworkers without further specification	0
9931	[9931] Forepersons, group leaders without further specification	0
9941	[9941] Persons doing community service without further specification	0
9951	[9951] Self-employed persons without further specification	0
9961	[9961] Consultancy, planning professionals without further specification	0
9971	[9971] Other employees without further specification	47
-1	[-1] No Answer	7
-2	[-2] Does not apply	29533
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: -1985
The variable is designed to provide information on the occupation of civil servant training

("Beamtenausbildung"). See the description of variable TRAINA$$ for more details on the construction and the values of the variable. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

fdt_f$$ – Data source FIELD, DEGREE, TRAIN [generic]

1	[1] Individual Questionnaire	0
2	[2] Gap Questionnaire (temporary drop-outs)	0
3	[3] Biographical Questionnaire	29935
4	[4] Various Sources	0
-1	[-1] No Answer	0
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: -1985
This is a flag variable which provides information on the data sources used for the construction of the variables FIELD$$, DEGREE$$, TRAINA$$, TRAINB$$, TRAINC$$ and TRAIND$$ (see the description of the respective variables for details). [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

bilztch$$ – Change in Education since last survey / last year [generic]

0	[0] Consistent educational information since last survey	21981
1	[1] Inconsistent educational information since last survey	9
2	[2] Inconsistent educational information since last year	0
-1	[-1] No Answer	0
-2	[-2] Does not apply	7945
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: -1985
This is a flag variable which identifies observations with inconsistent changes in the information on highest educational qualification compared to the previous observation or year.
Questions on highest educational attainment have been asked in the first survey and were only updated in subsequent years if the respondent reported a change. In the year 2000, every single SOEP participant was asked his highest level of educational attainment which produced a number of inconsistencies between the most recent information from 2000 and the generated information from previous years. These inconsistencies include both higher and lower educational attainment and are not just due to repeating the question about educational attainment in 2000. They also occur more generally, although to a lower degree, in the second survey wave of new samples when respondents to individual and life history questionnaires are asked to state their educational attainment. In both situations, respondents are not only asked annual questions about any changes in educational attainment since the

previous year, but are also asked to state their highest level of educational attainment.
In our view there is no means of unequivocally correcting for these inconsistencies. The flag variable helps researchers to identify observations with inconsistent answers to educational questions in the cross-sectional perspective. Researchers need to decide how to deal with these on a case-bycase basis depending on the research question at hand.
So far, we have not found evidence that respondents with a change in the year 2000 differed systematically from other respondents. One possible approach would be to exclude these individuals from the analysis when sample size allows. Alternatively, one could apply the information collected in 2000 to the prior years in which no changes were recorded between two years and test whether the results differ from those obtained when these individuals are left out. Since 2011, a Beta version of BIOEDU has also been made available, containing new data on consistent longitudinally tested educational transitions. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

bilztev$$ – Change in Education, total observed period [generic]

0	[0] Consistent educational information	20168
1	[1] Inconsistent educational decline	808
2	[2] Inconsistent educational increase	1697
3	[3] Inconsistent educational decline and increase	128
-1	[-1] No Answer	0
-2	[-2] Does not apply	7134
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: -1985
This flag variable identifies observations with at least one inconsistent change in the information given on individual highest educational qualification over the whole observation period. See the description of variable BILZTCH$$ for more details on the sources of these inconsistencies. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

11 Political and Religious Attitudes

picoriga16 – Party Identification in Country of Origin

1	[1] Yes	242
2	[2] No	4149
-1	[-1] No Answer	136
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	25408
-6	[-6] Version of questionnaire with modified filtering	0

picorigb16 – Party Identified with in Country of Origin

AF_.a_Afghan Nation		9
AF_.a_Afghanistan Republic Party		4
AF_.a_Islamic Society		3
AF_.a_Islamic Unity Party of Afghanistan		2
AF_.a_National-Islamic Front of Afghanistan		1
AF_.a_National Islamic Movement of Afghanistan		3
AF_.a_Peoples Islamic Movement of Afghanistan		2
AF_.a_Peoples Islamic Unity Party of Afghanistan		7
AL_.a_Red and Black Alliance		1
AL_PD_Democratic Party of Albania		4
AL_PS_Socialist Party of Albania		1
AM_HAK_Armenian National Congress		1
AM_OEK_Rule of Law		1
does not apply		4285
ER_.a_Eritrean Democratic Alliance		1
...	(21 rows omitted)	25560
RS-KM_.a_Serbian List		2
RS-KM_LDK_Democratic League of Kosovo		1
RS-KM_PDK_Democratic Party of Kosovo		1
RS_DS_Democratic Party		1
RU_ER_United Russia		1
RU_LDPR_Liberal Democratic Party of Russia		3
RU_SR_A Just Russia		1
SO_.a_Democratic Party of Somalia		1
SO_.a_United Somali Parliamentarians		1
SY_.a_Arab Socialist Ba'ath Party		15
SY_.a_Arab Socialist Union of Syria		1
SY_.a_Free Syrian Army		7
SY_.a_Syrian Communist Party		5
SY_.a_Syrian National Coalition		9
UA_.a_Party of Regions Opposition Bloc		1

picorigc16 – Party Family of PI in Country of Origin

0	[0] Not Categorized	57
1	[1] ECO_Ecological Parties	0
2	[2] COM_Socialist Parties	7
3	[3] SOC_Social Democratic Parties	6
4	[4] LIB_Liberal Parties	3
5	[5] CHR_Christian Democratic Parties	0
6	[6] CON_Conservative Parties	10
7	[7] NAT_Nationalist Parties	23
8	[8] AGR_Agrarian Parties	0
9	[9] ETH_Ethnic and Regional Parties	22
10	[10] SIP_Special Issue Party	0
11	[11] ISP_Islamic Parties	7
-1	[-1] No Answer	107
-2	[-2] Does not apply	4285

-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	25408
-6	[-6] Version of questionnaire with modified filtering	0

bgp_reli – Belong To Church Or Religious Community

0	[0] No Denomination	7813
11	[11] Catholic	5641
12	[12] Protestant	6471
13	[13] Christian Orthodox	850
14	[14] Other Christian Religion	562
20	[20] Islamic, not specified	409
21	[21] Shiite	337
22	[22] Sunni	2904
23	[23] Alevi	48
24	[24] Other isl. Religion	264
30	[30] Other Religious Denomination	522
-1	[-1] No Answer	471
-2	[-2] Does not apply	3643
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

The Variable bgp_reli summarizes the religious denomination of the respondents. Due to the fact that this question is part of the SOEP rotating scheme this variable is a summary of different points in time in regard of the different sub-samples. The information for the IAB-BAMF-SOEP Survey of Refugees (M3 and M4) can be identified in the Wave v33 whereas the information for all other respondents can be identified in Wave v32.
Please note, that the question may slightly vary between different sub-samples. Therefore, in some cases the specific religious denomination cannot be identified.

12 Information on the Interview

month$$ – Month Of Interview [generic]

1	[1] January	178
2	[2] February	7486
3	[3] March	5298
4	[4] April	3949
5	[5] May	3011
6	[6] June	1254
7	[7] July	1079
8	[8] August	2660
9	[9] September	1885
10	[10] October	941
11	[11] November	1337
12	[12] December	857
-1	[-1] No Answer	0

-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
Month of interview is generated using the answers to the individual questionnaire. Missing answers are filled in using data from the $hbrutto files. Interviews that took place in December and before the 20th of that month were recoded -3. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

mode$$ – Interview Method [generic]

100	[100] With Interviewer Assistance	0
110	[110] Oral Interview	1324
120	[120] Written Ques. Interviewer	3107
130	[130] Mixed Type	0
131	[131] Written Ques. No Interviewer	258
132	[132] Oral And Written	296
133	[133] Proxy	0
134	[134] Third Person Present	0
135	[135] No Third Person Present	0
140	[140] CAPI - Since 1998 (O)	20789
150	[150] Cawi Since 2014 (BE)	1547
200	[200] Telephone Assistance	0
210	[210] Written, By Mail	2390
220	[220] Telephone Interview	2
-1	[-1] No Answer	222
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0

Waves: all
The interview method is generated via the answers to the questions in the individual questionnaire. Missing answers are filled in from the $pbrutto files. [This information can be related to a specific variable and is not necessarily generic.]
For more information, contact: Peter Krause (Tel. +49-30-89789-690)