

DIW Berlin / SOEP (Ed.)

## Research Report

# SOEP-Core v38.1 - PGEN: Person-related status and generated variables

SOEP Survey Papers, No. 1313

### Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

*Suggested Citation:* DIW Berlin / SOEP (Ed.) (2024) : SOEP-Core v38.1 - PGEN: Person-related status and generated variables, SOEP Survey Papers, No. 1313, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/297845>

#### Standard-Nutzungsbedingungen:

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

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

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

#### Terms of use:

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

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

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



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

# 1313<sup>2024</sup>

**SOEP** Survey Papers  
Series D – Variable Descriptions and Coding

## SOEP-Core v38.1 – PGEN: Person-Related Status and Generated Variables

SOEP Group

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/soepsurveyspapers>

Editors:

Dr. Carina Cornesse, DIW Berlin and University of Bremen

Dr. Jan Goebel, DIW Berlin

Prof. Dr. Cornelia Kristen, University of Bamberg and DIW Berlin

Prof. Dr. Philipp Lersch, DIW Berlin and Humboldt-Universität zu Berlin

Prof. Dr. Carsten Schröder, DIW Berlin and Freie Universität Berlin

Prof. Dr. Jürgen Schupp, DIW Berlin and Freie Universität Berlin

Prof. Dr. Sabine Zinn, DIW Berlin and Humboldt-Universität zu Berlin

Please cite this paper as follows:

SOEP Group, 2024. SOEP-Core v38.1 – PGEN: Person-Related Status and Generated Variables. SOEP Survey Papers 1313: Series D – Variable Descriptions and Coding. Berlin: DIW Berlin/SOEP



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

© 2024 by SOEP

ISSN: 2193-5580 (online)

DIW Berlin

German Socio-Economic Panel (SOEP)

Mohrenstr. 58

10117 Berlin

Germany

[soeppapers@diw.de](mailto:soeppapers@diw.de)

# SOEP-Core v38.1 – PGEN: Person-Related Status and Generated Variables

SOEP Group

2024

## Inhaltsverzeichnis

<b>1</b>	<b>General Information</b>	<b>5</b>
<b>2</b>	<b>Identifiers</b>	<b>5</b>
	pid – Never Changing Person ID . . . . .	5
	hid – Current Wave HH Number . . . . .	5
	cid – Original Household Number, Case ID . . . . .	5
	syear – Survey Year . . . . .	5
<b>3</b>	<b>Own Nationality and Residential Status</b>	<b>6</b>
	pgnation – 1. Nationality . . . . .	6
	pgnationiso3n – 1. Nationality (ISO3N country code list) . . . . .	7
	pgstatus_asyl – Status quo, Application for Asylum . . . . .	8
	pgstatus_refu – Refugee residence status . . . . .	8
<b>4</b>	<b>Family Status and Partnership</b>	<b>8</b>
	pgpartz – Partner Indicator . . . . .	8
	pgpartnr – Partner Person Number . . . . .	9
	pgfamstd – Marital Status In Survey Year . . . . .	10
<b>5</b>	<b>Wages and Salary</b>	<b>11</b>
	pglabgro – Current Gross Labor Income in Euro . . . . .	11
	pgimpgro – Imputation flag for LABGROxx . . . . .	11
	pglabnet – Current Net Labor Income in Euro . . . . .	11
	pgimpnet – Imputation flag for LABNETxx . . . . .	12
	pgsndjob – Current Gross Secondary Income In Euro . . . . .	12
	pgimpsnd – Imputation Flag For SNDJOBxx . . . . .	13
	pgsndjob1 – Current Gross Income Secondary Job (1) In Euro . . . . .	13
	pgsndjob2 – Current Gross Income Secondary Job (2) In Euro . . . . .	14
	pgsndjob3 – Current Gross Income Secondary Job (3) In Euro . . . . .	15
	pgimpsnd1 – Imputation flag, Current Gross Income Secondary Job (1) . . . . .	16
	pgimpsnd2 – Imputation flag, Current Gross Income Secondary Job (2) . . . . .	16
	pgimpsnd3 – Imputation flag, Current Gross Income Secondary Job (3) . . . . .	17
<b>6</b>	<b>Current Employment Status</b>	<b>17</b>
	pgstib – Occupational Position . . . . .	17
	pgemplst – Employment Status . . . . .	18
	pglfs – Labor Force Status . . . . .	19
	pgjobch – Occupational Change . . . . .	20
	pgsndtyp1 – 1. secondary job: Occupational Position . . . . .	22
	pgsndtyp2 – 1. secondary job: Occupational Position . . . . .	23
	pgsndtyp3 – 3. secondary job: Occupational Position . . . . .	23
<b>7</b>	<b>Current Occupation</b>	<b>23</b>
	pgisco88 – Current Occupational Classification (ISCO-88 Com) . . . . .	24
	pgisco08 – Current Occupational Classification (ISCO-08) . . . . .	25
	pgkldb92 – Current Occupational Classification (KldB92) . . . . .	26
	pgkldb2010 – Current Occupational Classification (KldB2010) . . . . .	26
	pgisei08 – Last Reached Isei Value (International Socio-Economic Index) . . . . .	27
	pgmps08 – Last Reached Mps Value (Magnitude-Prestige-Skale, Wegener) . . . . .	28
	pgsiops08 – last reached SIOPS score (stand. internat. occ. prestige scale, Treiman) . . . . .	29

pggep08 – Last Reached Egp Value (Erikson, Goldthorpe, Portocarero) . . . . .	30
pgisei88 – Last Reached Isei Value (International Socio-Economic Index) . . . . .	31
pgmps92 – Last Reached Mps Value (Magnitude-Prestige-Skale, Wegener) . . . . .	32
pgsiops88 – last reached SIOPS score (stand. internat. occ. prestige scale, Treiman) . . . . .	33
pggep88 – Last Reached Egp Value (Erikson, Goldthorpe, Portocarero) . . . . .	33
pgerljob – Working In Occupation Trained For . . . . .	34
pgausb – Required Training For Job . . . . .	35
pgerwzeit – Length Of Time With Firm . . . . .	36
pgtatzeit – Actual Work Time Per Week . . . . .	37
pgvebzeit – Agreed Upon Work Time Per Week . . . . .	37
pguebstd – Overtime Per Week . . . . .	38
pgoeffd – Civil Service . . . . .	38
pgnace – Industry Occupation [ pbra] (NACE Rev. 1.1, Sector) . . . . .	39
pgnace2 – Industry Occupation [ pbra] (NACE Rev. 2, Sector) . . . . .	40
pgbetr – Size of the Company . . . . .	41
pgallbet – Core Category Size Of The Company . . . . .	42
<b>8 Last Occupation</b>	<b>42</b>
pgjobend – Reasons for occupational change . . . . .	42
<b>9 Employment History</b>	<b>44</b>
pgexpft – Working Experience Full-Time Employment . . . . .	44
pgexppt – Working Experience Part-Time Employment . . . . .	45
pgexpue – Unemployment Experience . . . . .	46
<b>10 Volunteering</b>	<b>47</b>
pghonor1 – First Voluntary Work . . . . .	47
pghonor2 – Second Voluntary Work . . . . .	48
pghonor3 – Third Voluntary Work . . . . .	48
<b>11 School, Higher and Vocational Education</b>	<b>48</b>
pgiscd97 – ISCED-1997-Classification . . . . .	48
pgiscd11 – ISCED-2011-Classification . . . . .	49
pgcasmin – CASMIN Classification . . . . .	50
pgbilzeit – Amount Of Education Or Training In Years . . . . .	50
pgpsbil – School-Leaving Degree . . . . .	51
pgpbil01 – Vocational Degree Received . . . . .	52
pgpbil02 – College Degree . . . . .	52
pgpbil03 – No Vocational Degree . . . . .	53
pgpsbilo – School-Leaving Degree East Germany . . . . .	53
pgpbilo – Vocational Degree Received East Germany . . . . .	54
pgpsbila – School-Leaving Degree Outside Germany . . . . .	54
pgpbila – Vocational Degree Outside Germany . . . . .	55
pgfield – Field of tertiary education . . . . .	55
pgdegree – Type of tertiary degree . . . . .	57
pgtraina – Apprenticeship - two-digit occupation KldB92 . . . . .	59
pgtrainb – Vocational school - two-digit occupation KldB92 . . . . .	61
pgtrainc – Higher vocational school - two-digit occupation KldB92 . . . . .	62
pgtraind – Civil servant training - two-digit occupation KldB92 . . . . .	62
pgfdt_f – Data source FIELD, DEGREE, TRAIN . . . . .	63
pgbiltch – Change in Education since last survey / last year . . . . .	64

pgbilztev – Change in Education, total observed period . . . . .	64
<b>12 Political Attitudes</b>	<b>65</b>
pgpicoriga – Party Identification in Country of Origin . . . . .	65
pgpicorigb – Party Identified with in Country of Origin . . . . .	65
pgpicorigc – Party Family of PI in Country of Origin . . . . .	66
<b>13 Information on the Interview</b>	<b>66</b>
pgmonth – Month Of Interview . . . . .	66
pgpiyear – Personal Related Survey Year . . . . .	67
pgmode – Interview Method . . . . .	67
inputdataset – Source dataset . . . . .	68

## 1 General Information

The PGEN-files contain user friendly data on the individual level which are consolidated from different sources in long format. The plausibility is in many respects longitudinally validated, therefore the data here are in most situations superior compared to the data in PL or \$P.

The file contains one row for each person (PID) and surveyyear (SYEAR) (each combination of PID and SYEAR is unique) with a completed personal or youth questionnaire. These are the persons where NETTO in PPFADL has the values 10–17 or 19 which is equivalent for values 1 and 5 in NETT1.

The same content can be found in the wave specific \$PGEN files. The variable names are similar and differ especially in prefixes and suffixes.

In 2017 the file contains information about two special groups of first-time respondent refugee teenagers. First group of 143 respondents (netto 29) is analogous to 2016, part of the household questionnaire and the data are stored in the child dataset. The second group contains 108 individuals of age 16-17 from refugee sample, who received a special questionnaire. These respondents are part of youth dataset JUGENDL and can be identified by the instrument indicator value 73.

From version v34 on, PGAUTONO is not a part of data delivery.

From version v32 on, PGERWTYP is not a part of data delivery. The information related to PGERWTYP categories can be found in the variable(s) PGJOBCH, PGEMPLST, PGLFS.

## 2 Identifiers

### **pid** – Never Changing Person ID

---

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

### **hid** – Current Wave HH Number

---

This identifier groups all persons into households at the time of the most recent wave.

### **cid** – Original Houshold Number, Case ID

---

The identifier of the household, when it is sampled and selected for interviewing for the first time. The CID 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.

### **syear** – Survey Year

---

1984	12245
1985	11090
1986	10646
1987	10516
1988	10023
1989	9710
1990	13972
1991	13669



1992	13397
1993	13179
1994	13417
1995	13768
1996	13511
1997	13283
1998	14670
... (8 rows omitted)	173304
2007	21232
2008	19945
2009	21035
2010	27124
2011	29264
2012	28520
2013	31523
2014	28042
2015	27743
2016	29870
2017	33265
2018	30997
2019	30397
2020	31083
2021	20770

### 3 Own Nationality and Residential Status

#### p<sub>gnation</sub> – 1. Nationality

---

1	Germany	640519
2	Turkey	23509
3	Ex-Yugoslavia	7448
4	Greece	8243
5	Italy	12275
6	Spain	4658
7	Ex-GDR (Country Of Origin Only)	0
10	Austria	1581
11	France	773
12	Benelux	119
13	Denmark	189
14	Great Britain	628
15	Sweden	103
16	Norway	37
17	Finland	104
...	(171 rows omitted)	49889
189	Central African Republic	0
190	Djibouti	0
191	Gabon	0
192	Comoros	0
193	Katar	0
194	Sahara	0

195	Ingushetia	8
196	Kosovo	250
222	Eastern Europe	0
333	Other Unspecified Foreign Country	0
999	ethnic minorities (e.g. Yazidi, Roma)	2
-1	No answer	99
-2	Does not apply	6
-7	Only available in less restricted edition	0
-8		20770

*Waves: –2020*

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 PGNATION thus offers a variable on nationality for all subsamples.

*For more information, contact: Peter Krause (Tel. +49-30-89789-690)*

### **pgnationiso3n** – 1. Nationality (ISO3N country code list)

---

4	Afghanistan	202
8	Albania	49
10	Antarctica	0
12	Algeria	2
16	American Samoa	0
20	Andorra	0
24	Angola	1
28	Antigua and Barbuda	0
31	Azerbaijan	16
32	Argentina	3
36	Australia	5
40	Austria	43
44	Bahamas	0
48	Bahrain	0
50	Bangladesh	12
...	(229 rows omitted)	20287
862	Venezuela	3
876	Wallis and Futuna	0
882	Samoa	3
887	Yemen	5
894	Zambia	0
900	Kosovo	57
997	stateless	62
-1	No answer	18
-2	Does not apply	2
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0

-7	Only available in less restricted edition	0
-8	Question this year not part of survey	750440

Waves: 2021–

### pgstatus\_asyl – Status quo, Application for Asylum

---

1	[ de] Schutzstatus zuerkannt (Anerkannter Fluechtling, Asylberechtigt, anderer Schutzstatus)	4999
2	Temporary Suspension of Deportation	513
3	Request to Leave Germany	343
4	Decision Regarding Residential Status Still Open	1948
-1	No answer	103
-2	Does not apply	27316
-5	Not Included In Questionnaire Version	27913
-7	Only available in less restricted edition	0
-8	Question this year not part of survey	708075

This variable provides data on the current status of the application for asylum. The data are obtained by asking respondents if their application for asylum has been approved and if an official decision regarding the application for asylum has been made. Please also see pgstatus\_refu and the variable arefback in the dataset ppfadl.

For more information, contact: Jana Nebelin (Tel. +49 30 89789-319; jnebelin@diw.de)

### pgstatus\_refu – Refugee residence status

---

1	In Proceedings	2459
2	[ de] Schutzstatus zuerkannt (Asyl, Fluechtlingseigenschaft, Niederlassungserlaubnis, human. Aufnahme, sonst. human. Aufnahme)	6381
3	Temporary Suspension of Deportation	559
4	Other	467
-1	No answer	170
-2	Does not apply	25186
-5	Not Included In Questionnaire Version	27913
-7	Only available in less restricted edition	0
-8	Question this year not part of survey	708075

This variable provides data on the current status of the refugee. The data are obtained by asking respondents for their current residence title. Please also see pgstatus\_asyl and the variable arefback in the dataset ppfadl.

For more information, contact: Jana Nebelin (Tel. +49 30 89789-319; jnebelin@diw.de)

## 4 Family Status and Partnership

### pgpartz – Partner Indicator

---

0	No partner	243126
1	Spouse, registered partner	460369
2	Partner	64284

3	Probably spouse, registered partner	1078
4	Probably partner	1746
5	not clear	607
-1	No answer	0
-2	Does not apply	0
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

*Waves:* all

The variable PGPARTZ generated in the context of the partner identifier (PGPARTNR) 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 5. 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.

*For more information, contact:* Claudia Saalbach (Tel. +49-30-89789-320, csaalbach@diw.de)

### pgpartnr – Partner Person Number

-1	No answer	0
-2	Does not apply	243726
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

*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 PGPARTZ is coded 0, 3, 4, 5 in these cases. In couples, PGPARTNR is the value of the unchanging person ID number (=PID) 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 (PLK0001\_\* in PL), the household matrix reported by the head of household at the beginning of the interview (STELL\_H in PBRUTTO), 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. Since 2021, the relmatrix data set is also used to create partnership links. In unclear cases, due to temporal non-response for instance, we also consider

longitudinal information from previous and prospective waves. Moreover, PGPARTNR 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.

*For more information, contact:* Claudia Saalbach (Tel. +49-30-89789-320, csaalbach@diw.de)

### pgfamstd – Marital Status In Survey Year

---

1	Married	456740
2	Married, But Separated	16615
3	Single	190892
4	Divorced	54047
5	[ de] Verwitwet / Lebenspartner/-in aus eingetragener gleichgeschlechtlicher Partnerschaft verstorben	43566
6	husband/wife abroad	2034
7	Registered Same-Sex Partnership, Living Together	899
8	Registered Same-Sex Partnership, Living Apart	147
-1	No answer	3099
-2	Does not apply	0
-3	Not valid	3164
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	7
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	0

*Waves:* all

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 file BIOBIRTH might 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 file, although value labels differ between PGFAMSTD and SPELLTYP in BIOMARSY. 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 PGPARTZ supplied here in dataset PGEN as well might not match the information provided in PGFAMSTD in its entirety.

*For more information, contact:* Jascha Dräger (Tel. +49 30 89789-367, jdraeger@diw.de)

## 5 Wages and Salary

### pglabgro – Current Gross Labor Income in Euro

---

-2	Does not apply	314165
-5	Not Included In Questionnaire Version	5286
-7	Only available in less restricted edition	0

*Waves:* all

The variable PGLABGRO represents the imputed gross labor income in the previous month generated for all SOEP respondents who are employed in a main job in each respective wave. Interviewees should consider overtime payments but no irregular one-time payments such as holiday or bonuses. 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 (PGIMPGRO).

The corresponding variables in file PL are PLC0013\_V1 and PLC0013\_V2. The original variables coming from the \$P files 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, bhp\_97\_01.

*For more information, contact:* Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

### pgimpgro – Imputation flag for LABGROxx

---

0	Not imputed	404336
1	Imputed	47423
-1	No answer	0
-2	Does not apply	314165
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	5286
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

*Waves:* all

The variable PGIMPGRO designates imputations of item non-response in the variable PGLABGRO (current gross labor income). The value -2 (does not apply) can be interpreted as “not working”.

*For more information, contact:* Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

### pglabnet – Current Net Labor Income in Euro

---

-2	Does not apply	314166
-5	Not Included In Questionnaire Version	5286
-7	Only available in less restricted edition	0

*Waves:* all

The variable PGLABNET represents the imputed net labor income in the previous month in a main job of all persons in SOEP working in the respective wave. Interviewees should consider overtime payments but no irregular one-time payments such as holiday or bonuses. Income details are consistently provided in euros for all waves. The imputation of item non-response takes place in a two-stage procedure: first, with the “Row-and- Column” method of Little und Su (1989) using individual longitudinal data as well as cross-sectional trend data (cf. Joachim R. Frick and Markus M. Grabka (2005): Item-Non-Response on Income Questions in Panel surveys: Incidence, Imputation and the Impact on the Income Distribution. Allgemeines Statistisches Archiv (ASTA) 89, 49-61). Alternatively, if no individual longitudinal information is available, we base the imputation on a regression using different Mincer covariates, also taking into account current gross labor income. If both types of income information are lacking, first we impute current gross labor income and then current net labor income.

Imputed values are flagged (PGIMPNET).

The corresponding variables in file PL are PLC0014\_V1 and PLC0014\_V2. The original variables coming from the \$P files 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, bhp\_97\_02.

*For more information, contact:* Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

### pgimpnet – Imputation flag for LABNETxx

---

0	Not imputed	417999
1	Imputed	33759
-1	No answer	0
-2	Does not apply	314166
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	5286
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

*Waves:* all

The variable PGIMPNET designates imputations of item non-response in the variable PGLABNET (current net labor income). The value -2 (does not apply) can be interpreted as “not working”.

*For more information, contact:* Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

### pgsndjob – Current Gross Secondary Income In Euro

---

-2	Does not apply	447641
-5	Not Included In Questionnaire Version	292615
-7	Only available in less restricted edition	0

Waves: 1995-2016

The variable PGSNDJOB 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 PGIMPSND in year 2013 was set to -1 and PGSNDJOB had been imputed. Imputed values are flagged (PGIMPSND).

The corresponding variables in file PL are PLC0062\_V1 and PLC0062\_V2. The original variables coming from the \$P files are: lp7702, mp5802, np5802, op4902, pp6602, qp6302, rp6602, sp6602, tp8402, up67, vp79, wp70, xp81, yp76, zp78, bap69, bbp79, bcp67, bdp85, bep64, bfp113, bgp97.

For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

### pgimpsnd – Imputation Flag For SNDJOBxx

---

0	Not imputed	27842
1	Imputed	3112
-2	Does not apply	447641
-5	Not Included In Questionnaire Version	292615
-7	Only available in less restricted edition	0

Waves: 1995-2016

The variable PGIMPSND designates imputations of itemnonresponse in the variable PGSNDJOB\$\$ (current gross labor income from second job). The value -2 (does not apply) can be interpreted as “not working”.

For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

### pgsndjob1 – Current Gross Income Secondary Job (1) In Euro

---

0	553
1	27
2	4
3	3
4	4
5	9
7	2
8	3
9	4
10	35
12	9



13		1
14		2
15		23
16		9
...	(639 rows omitted)	8053
9000		3
9800		1
10000		1
10173		1
10293		1
12000		1
12500		2
13000		1
15000		2
43000		1
45000		1
100000		1
-2	Does not apply	114256
-5	Not Included In Questionnaire Version	648197
-7	Only available in less restricted edition	0

*Waves:* all

The variable PGSNDJOB1 represents the imputed gross labor income from a first side job. This information is collected for the first time since 2017. 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. Alternatively, if no individual longitudinal information is available, we base the imputation on a regression using different Mincer covariates. Imputed values are flagged (PGIMPSNDJOB1).

*For more information, contact:* Markus Grabka (Tel. +49-30-89789-339 / [mgrabka@diw.de](mailto:mgrabka@diw.de))

### **pgsndjob2** – Current Gross Income Secondary Job (2) In Euro

---

0		157
1		3
2		1
3		1
4		1
5		5
6		2
7		4
8		1
9		1
10		11
12		3
14		3
15		9
16		4
...	(240 rows omitted)	1151
2500		2
2950		1

3000	4
3028	1
3900	1
4000	2
4150	1
4395	1
4500	1
6815	1
8000	1
20000	1
-2 Does not apply	121639
-5 Not Included In Questionnaire Version	648197
-7 Only available in less restricted edition	0

Waves: 2017–

The variable PGSNDJOB2 represents the imputed gross labor income from a first side job. This information is collected for the first time since 2017. 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. Alternatively, if no individual longitudinal information is available, we base the imputation on a regression using different Mincer covariates. Imputed values are flagged (PGIMPSNDJOB2).

For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

### pgsndjob3 – Current Gross Income Secondary Job (3) In Euro

---

0	41
1	1
2	2
5	1
6	1
8	1
9	1
10	4
11	1
12	1
13	1
15	4
17	1
19	1
20	7
... (86 rows omitted)	263
1760	1
1800	1
1954	1
2000	2
2300	1
4000	2
5000	2
10000	1
13149	1

25000		1
172944		1
432566		1
-2	Does not apply	122667
-5	Not Included In Questionnaire Version	648197
-7	Only available in less restricted edition	0

*Waves: 2017–*

The variable PGSNDJOB3 represents the imputed gross labor income from a first side job. This information is collected for the first time since 2017. 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. Alternatively, if no individual longitudinal information is available, we base the imputation on a regression using different Mincer covariates. Imputed values are flagged (PGIMPSNDJOB3).

*For more information, contact:* Markus Grabka (Tel. +49-30-89789-339 / [mgrabka@diw.de](mailto:mgrabka@diw.de))

### **pgimpsnd1** – Imputation flag, Current Gross Income Secondary Job (1)

---

0	Not imputed	8101
1	Imputed	656
-1	No answer	0
-2	Does not apply	114256
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	648197
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	0

*Waves: 2017–*

The variable PGIMSND1 designates imputations of item non-response in the variable PGSNDJOB1 (current gross labor income from a side job).

*For more information, contact:* Markus Grabka (Tel. +49-30-89789-339 / [mgrabka@diw.de](mailto:mgrabka@diw.de))

### **pgimpsnd2** – Imputation flag, Current Gross Income Secondary Job (2)

---

0	Not imputed	1154
1	Imputed	220
-1	No answer	0
-2	Does not apply	121639
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	648197
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	0

*Waves: 2017–*

The variable PGIMSND2 designates imputations of item non-response in the variable PGSNDJOB2 (current gross labor income from a side job).

*For more information, contact:* Markus Grabka (Tel. +49-30-89789-339 / [mgrabka@diw.de](mailto:mgrabka@diw.de))

**pgimpsnd3** – Imputation flag, Current Gross Income Secondary Job (3)

0	Not imputed	263
1	Imputed	83
-1	No answer	0
-2	Does not apply	122667
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	648197
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	0

Waves: 2017–

The variable PGIMSND3 designates imputations of item non-response in the variable PGSND-JOB3 (current gross labor income from a side job).

For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

**6 Current Employment Status****pgstib** – Occupational Position

10	Not employed (NE): without further information	81707
11	NE: In education, further education, apprenticeship	36795
12	NE: registered unemployed	47978
13	NE: Pensioner	153848
15	NE: Military, Community Service	2200
110	Apprentice (until 1999), trainee (1990 East)	6985
120	Apprentice / trainee in industry or technology (since 2000)	8853
130	Apprentice / trainee in trade or commerce (since 2000)	4893
140	Student trainee, intern	2410
150	Aspirant	3
210	Untrained Worker	22957
220	Semi-Trained Worker	45850
230	Trained Worker	49158
240	Foreman, group leader	6626
250	Foreman	2924
...	(21 rows omitted)	105745
530	Salaried employee, skilled	98656
540	Salaried employee, highly skilled or managerial duties	55210
550	Salaried employee with extensive managerial duties	6810
560	Managing partner or comparable employee in own company (since 2019)	579
610	Civil servant, lower level	979
620	Civil servant, middle level	7387
630	Civil servant, upper level	12100
640	Civil servant, executive level	7749
-1	Employed without information on occupational position	2385
-2	Does not apply	336
-3	Not valid	0

-4	Inadmissible multiple response	0
-5	Not Included In Questionnaire Version	87
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0

*Waves:* all

The variable represents a compilation of all relevant information on current occupational position. It is generated by combining the main input variables p1b0064\_v# (occupational group employee), p1b0058 (occupational group industrial worker), p1b0057\_# (occupational group self-employed), p1b0022-v# (employment status), p1c0232\_v# (pension), and p1b0021 (registered unemployed) from the pl dataset. The following list shows the prioritization of data when there are multiple conflicting sources of information (increasing priority from top to bottom):

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 PGSTIB, non-working respondents 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 file 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 2014, it is possible that self-employed people (410–440) leave missing the subsequent question about the number of employees or co-workers used for the categorization. These observations are then assigned to their respective primary categories (410, 420 and 430) which were otherwise valid from 1984 until 1999.

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

### pgemplst – Employment Status

1	Full-Time Employment	307830
2	Regular Part-Time Employment	85472
3	Vocational Training	22493
4	Marginal, Irregular Part-Time Employment	34585
5	Not Employed	319580

6	Sheltered workshop (since 1998)	700
7	In short-time work (since 2021)	200
-1	No answer	134
-2	Does not apply	65
-3	Not valid	2
-4	Inadmissible multiple response	22
-5	Not included in this version of the questionnaire	127
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0

*Waves:* all

This variable is generated from the annual question on current employment status represented in the p1b0022\_v# variables from the pl dataset. PGEMPLST 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 people 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 people in sheltered employment.

Starting in 2021, respondents were given an additional response option to indicate whether they were currently on short-time work. As a result, a new category with the value (7) for employees in short-time work was introduced in PGEMPLST starting from wave 38.

PGEMPLST supplements the variable PGLFS, which differentiates among individuals who are not employed.

The information about newly included teenager refugee sample special groups in 2017 is taken from child and youth datasets. In case the teenagers do not attend school, they were asked about their alternative status. Those teenagers who attend school, are in vocational preparation year, in a vocational training, or in vocational school were assigned value (3). Those waiting for the start of the school term or vocational training, are categorized as (1). In case the teenagers do not attend school and are working the respondents were classified to category (3). This information also applies for first-time teenage respondents in 2016 (who were previously treated as missings).

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

### pglfs – Labor Force Status

1	Non-working (NW): without further information	92102
2	NW: age 65 and older	112852
3	NW: and currently in training/education	32839
4	NW: on parental leave (since 1991)	14087
5	NW: in military/community service	1005
6	NW: and registered unemployed	42223
8	NW: but occasional secondary job (1985-2016)	8581
9	NW: but paid work in past 7 days (since 1999)	4817
10	NW: but regular secondary job (1985-2016)	7540
11	Working	446496
12	Working, but inactive within past 7 days (since 2000)	6809
13	NW: but paid secondary job (since 2017)	1765

-1	No answer	1
-2	Does not apply	93
-3	Not valid	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	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. It differentiates between “working” (Code 11–12) and “non-working” (Code 1–10,13) 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. PGLFS is generated mainly from plb0073\_v variables, (employment status) from the pl dataset. The number of values assigned is based on a number of differentiated answer categories. It is designed to provide consistent longitudinal data on labor force participation across all waves.

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 priority from top to bottom):

value	label
1	non-working without further information
2	non-working, and age 65 and older
3	non-working, and currently in training/education
6	non-working, and registered unemployed
4	non-working, on maternity leave
5	non-working, in military/community service
8	non-working, but occasional secondary job (1985-2016)
10	non-working, but regular secondary job (1985-2016)
9	non-working, but paid work in past 7 days
13	non-working, but some secondary job
11	working
12	working, but inactive in the past 7 days

Since 2016, the questionnaire for refugees contains a new labor market status category “vocational internship (betriebliches Praktikum)”, which is integrated in the category (3) “non-working, and currently in training/education” of PGLFS.

Since 2017 there is no differentiation between regular and occasional secondary job. Category (13) is therefore created to capture the respondents who are primarily not part of the labor force but either regularly or occasionally exercise one or more secondary activities which are either remunerated or not voluntary.

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

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

## pgjobch – Occupational Change

---

1	Not Employed	317667
2	Employed No Change	367549
3	Employed No Info If Change	3826
4	Employed With Change	73931
5	First Time Employed	7971
-1	No answer	134
-2	Does not apply	107
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	25
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	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.

PGJOBCH is generated based on the central filter variable, which indicates whether a respondent 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. The main input variables are therefore: p1b0031\_v1 (No change of professional situation [1985-1993]), p1b0031\_v2 (Job change in the previous year [1994 - ...]), and p1b0022\_v# (employment status) from the pl dataset.

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 their first job.

The variable is also designed to provide consistent longitudinal information on job changes. The PGJOBCH variable is generated by correcting the original job change information in various ways:

1. We check whether the job changes stated by a respondent in two consecutive 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 PGJOBCH 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 PGJOBCH for the migration sample in 2015 was as following:

1. For follow-up respondents generation as in the case of a uniform questionnaire for all respondents was 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 applied.

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 PGJOBCH for these samples corresponds to the rule of thumb is as for the first-time respondents 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.

2017 generation of PGJOBCH for the newly added immigration sample [psample = 19] is analogous to that of 2016. As of refugee teenagers, those attending school or who are currently in vocational preparation year (Berufsvorbereitungsjahr), in vocational training, in a vocational integration course (ber. Eingliederungsjahrgang), or in vocational school (berufsbildende Schule) were assigned value (1). Those working were assigned jobch value (3).

*For more information, contact: Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)*

### pgsndtyp1 - 1. secondary job: Occupational Position

1	Worker/Employee (also mini job)	4445
2	Self-Employed, Free-Lance Professional	2934
3	Help In Family Business	578
4	Other occupational position	11090
-1	No answer	3772
-2	Does not apply	100194
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	648197
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	0

Waves: 2017–

The variable PGSNDTYP1 gives information about the current employment status in a first side job. This variable is a direct copy of the original variable in \$P.

For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

### pgsndtyp2 – 1. secondary job: Occupational Position

---

1	Worker/Employee (also mini job)	391
2	Self-Employed, Free-Lance Professional	609
3	Help In Family Business	119
4	Other occupational position	3736
-1	No answer	1447
-2	Does not apply	116711
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	648197
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	0

Waves: 2017–

The variable PGSNDTYP2 gives information about the current employment status in a second side job. This variable is a direct copy of the original variable in \$P.

For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

### pgsndtyp3 – 3. secondary job: Occupational Position

---

1	Worker/Employee (also mini job)	62
2	Self-Employed, Free-Lance Professional	166
3	Help In Family Business	32
4	Other occupational position	1194
-1	No answer	493
-2	Does not apply	121066
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	648197
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	0

Waves: 2017–

The variable PGSNDTYP3 gives information about the current employment status in a third side job. This variable is a direct copy of the original variable in \$P.

For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@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 PGJOBCH-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 necessarily consistent to the current employment status, because they are based on different sources of information.

### pgisco88 – Current Occupational Classification (ISCO-88 Com)

0	Soldiers	0
100	Soldiers	1559
1000	LEGISLATORS, SENIOR OFFICIALS AND MANAGERS	0
1100	Legislators, Senior Officials and Managers	0
1110	Legislators, Senior Officials and Managers	136
1140	Senior Officials of Special-Interest Organisations	50
1141	Senior Officials of Political Party Organisations	9
1142	Senior officials of employers', workers' and other economic-interest organisations	257
1143	Senior Officials of Humanitarian and Other Special-Interest Organisations	4
1200	Corporate Managers	1312
1210	Directors and Chief Executives	2891
1220	Production and Operations Managers	0
1221	Production and operations managers in agriculture, hunting, forestry and fishing	18
1222	Production and Operations Managers in Manufacturing	1662
1223	Production and Operations Managers in Construction	0
...	(470 rows omitted)	347590
9300	Labourers in mining, construction, manufacturing and transport	0
9310	Mining and Construction Labourers	0
9311	Mining and Quarrying Labourers	49
9312	Construction and maintenance labourers: roads, dams and similar constructions	31
9313	Building Construction Laborer	1182
9320	Manufacturing Laborer	6706
9330	Transport Lab., Freight Handler	3206
-1	No answer	2139
-2	Does not apply	289162
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	113247

Waves: 1984-2017

Current occupation coded as ISCO-88 COM (ISCO-88 version for European Union purposes). Respondents answer the question in their own words. If they don't report a jobchange, information from the previous year is used, cf. introduction to section in PDF documentation for this file. Hartmann and Schütz (2002) provide detailed information on the conducted occupational coding.

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](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)

### pgisco08 – Current Occupational Classification (ISCO-08)

0	Armed forces occupations	0
110	Commissioned armed forces officers	75
210	Non-commissioned armed forces officers	35
310	Armed forces occupations, other ranks	299
1000	Managers	0
1100	Chief executives, senior officials and legislators	0
1110	Legislators and senior officials	0
1111	Legislators	56
1112	Senior government officials	241
1113	Traditional chiefs and heads of village	0
1114	Senior officials of special-interest organizations	93
1120	Managing directors and chief executives	2742
1200	Administrative and commercial managers	0
1210	Business services and administration managers	1
1211	Finance managers	226
...	(565 rows omitted)	139728
9613	Sweepers and related labourers	33
9620	Other elementary workers	0
9621	Messengers, package deliverers and luggage porters	511
9622	Odd job persons	2
9623	Meter readers and vending-machine collectors	12
9624	Water and firewood collectors	0
9629	Elementary workers not elsewhere classified	170
-1	No answer / don't know	1430
-2	Does not apply	118036
-3	Implausible value	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	507520

Waves: 2013–

Current occupation coded as ISCO-08. Respondents answer the question in their own words. If they don't report a jobchange, information from the previous year is used, cf. introduction to section in PDF documentation for this file. Tschersich and Schütz (2014, SOEP Survey Paper 470) provide detailed information on occupational coding.

For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

**pgkldb92** – Current Occupational Classification (KldB92)

110	Farmers, general	1813
111	Fruit and vegetable farmers (non-horticultural)	5
112	Arable farmers (special, permanent crops)	9
113	Livestock farmers and pasture farmers	3
114	Seed, crop producers, propagators (non-horticultural)	8
115	Crop protectors	10
116	Farmers and landlords	32
118	Farmers and wine growers	12
120	Wine growers, general	27
121	Vine propagators	0
129	Other wine growers	0
130	Agricultural workers, general	105
131	Agricultural supervisors	7
132	Agricultural machinery drivers	58
133	Vineyard workers	65
...	(2263 rows omitted)	363511
9911	Specialized professionals without further specification	95
9921	Homeworkers without further specification	151
9931	Forepersons, group leaders without further specification	355
9941	Persons doing community service without further specification	9
9951	Self-employed persons without further specification	366
9961	Consultancy, planning professionals without further specification	213
9971	Other employees without further specification	1601
-1	No answer	717
-2	Does not apply	288791
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	113247

Waves: 1984-2017

Current occupation coded as KldB92. Respondents answer the question in their own words. If they don't report a jobchange, information from the previous year is used, cf. introduction to section in PDF documentation for this file. Hartmann and Schütz (2002) provide detailed information on occupational coding.

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](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](mailto:kwenzig@diw.de))

**pgkldb2010** – Current Occupational Classification (KldB2010)

1104	Commissioned officers	71
1203	Senior non-commissioned officers and higher	35
1302	Junior non-commissioned officers	15

1402	Armed forces personnel in other ranks	298
11101	Occupations in farming (without specialisation)-unskilled/semiskilled tasks	142
11102	Occupations in farming (without specialisation)-skilled tasks	483
11103	Occupations in farming (without specialisation)-complex tasks	12
11104	Occupations in farming (without specialisation)-highly complex tasks	61
11113	Technical occupations in farming-complex tasks	4
11114	Technical occupations in farming-highly complex tasks	0
11123	Agricultural experts-complex tasks	0
11124	Agricultural experts-high complex tasks	5
11132	Technical laboratory occupations in agriculture-skilled tasks	1
11133	Technical laboratory occupations in agriculture-complex tasks	0
11182	Occupations in farming (with specialisation, not elsewhere classified)-skilled tasks	6
...	(1264 rows omitted)	142850
94693	Supervisors in stage, costume and prop design	2
94704	Occupations in museums (without specialisation)-highly complex tasks	29
94712	Technical occupations in museums and exhibitions-skilled tasks	10
94713	Technical occupations in museums and exhibitions-complex tasks	1
94714	Technical occupations in museums and exhibitions-highly complex tasks	1
94724	Art experts-highly complex tasks	0
94794	Managers in museum	4
-1	No answer / don't know	1597
-2	Does not apply	118063
-3	Implausible value	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	507520

Waves: 2013–

Current occupation coded as KldB2010. Respondents answer the question in their own words. If they don't report a jobchange, information from the previous year is used, cf. introduction to section in PDF documentation for this file. Tschersich and Schütz (2014, SOEP Survey Paper 471) provide detailed information on occupational coding.

For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

### pgisei08 – Last Reached Isei Value (International Socio-Economic Index)

11.5600004196167	19
11.7399997711182	635
12.0100002288818	57
12.3400001525879	7
13.2399997711182	23
13.3400001525879	24
13.3500003814697	3
13.7200002670288	29
13.8699998855591	54
14.210000038147	5597
14.3900003433228	217
14.5699996948242	404

14.6400003433228		47
14.8199996948242		14
15.3500003814697		627
...	(338 rows omitted)	153959
85.4100036621094		1452
85.8499984741211		560
86.7200012207031		617
86.8099975585938		33
88.3099975585938		255
88.6999969482422		736
88.9599990844727		104
-1	No answer	1010
-2	Does not apply	97207
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question this year not part of survey	507520

*Waves: 2013–*

This variable reflects the (last reached) Standard International Socio-Economic Index of Occupational Status for all employed persons. The ISEI Index was developed in 1992 by Ganzeboom, De Graaf, Treiman, and De Leew based on information about income, education, and occupation. Technically, ISEI was created by scaling the ISCO classification. 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-08 code of the current occupation using the Stata ado derivescores by Daniel Bela and Knut Wenzig which itself is based on SPSS algorithms from Ganzeboom and Treimann (2010).

Also available: occupational prestige scores (SIOPS, MPS) and occupational class (EGP).

Ganzeboom, Harry B.G.; Treiman, Donald J., "International Stratification and Mobility File: Conversion Tools." Amsterdam: Department of Social Research Methodology, <http://www.harryganzeboom.nl/ismf/index.html> <2010-01-12>.

For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, [kwenzig@diw.de](mailto:kwenzig@diw.de))

### pgmps08 – Last Reached Mps Value (Magnitude-Prestige-Skale, Wegener)

20	287
23.8999996185303	724
24.7000007629395	656
26.7000007629395	438
26.8999996185303	2977
28.6000003814697	54
30	7585
30.2999992370605	228
31.2000007629395	1121
31.6000003814697	27
31.7999992370605	1341
31.8999996185303	699



32.4000015258789		1122
34.7000007629395		24
35.5999984741211		70
...	(163 rows omitted)	134945
159.800003051758		1475
160.300003051758		345
160.5		912
170.899993896484		617
173.300003051758		368
179.600006103516		1497
186.800003051758		104
-1	No answer	0
-2	Does not apply	106074
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question this year not part of survey	507520

Waves: 2013–

This variable gives the occupational prestige score developed by Wegener (1988) for all employed persons. Like the SIOPS prestige score, Wegener's prestige scale measures a person's occupational prestige and was developed especially for use in the Federal Republic of Germany.

MPS08 is assigned based on ISCO-88 (using Christoph 2005), which is based on ISCO-08 (using Ganzeboom, Treiman 2010).

Also available: occupational prestige scores (SIOPS, ISEI) and occupational class (EGP).

Christoph, Bernhard: 2005: *Zur Messung des Berufsprestiges: Aktualisierung der Magnitude-Prestigeskala auf die Berufsklassifikation ISCO88*, <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-207543> –

Ganzeboom, Harry B.G.; Treiman, Donald J., "International Stratification and Mobility File: Conversion Tools." Amsterdam: Department of Social Research Methodology, <http://www.harryganzeboom.nl/ismf/index.html>

<2010-01-12>.

For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, [kwenzig@diw.de](mailto:kwenzig@diw.de))

**pgsiops08** – last reached SIOPS score (stand. internat. occ. prestige scale, Treiman)

12	29
13	271
15	627
16	274
17	17
18.0200004577637	12
18.9599990844727	3
20	6694
20.0300006866455	418
20.0900001525879	2
20.2600002288818	35
20.2999992370605	2069
20.3899993896484	767



20.7600002288818		8
21		478
...	(216 rows omitted)	148849
73.0999984741211		617
73.5100021362305		985
75		205
75.6800003051758		60
76.1100006103516		104
78.0100021362305		1497
78.1600036621094		1452
-1	No answer	1010
-2	Does not apply	97207
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question this year not part of survey	507520

Waves: 2013–

This variable gives the occupational prestige score index for all employed persons. It is derived from the ISCO-08 code of the current occupation using using the Stata ado derivescores by Daniel Bela and Knut Wenzig which itself is based on SPSS algorithms from Ganzeboom and Treiman (2010).

Please also see occupational prestige scores (MPS), occupational status (ISEI), and occupational class (EGP).

Ganzeboom, Harry B.G.; Treiman, Donald J., "International Stratification and Mobility File: Conversion Tools." Amsterdam: Department of Social Research Methodology, <http://www.harryganzeboom.nl/ismf/index<2010-01-12>>.

For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, [kwenzig@diw.de](mailto:kwenzig@diw.de))

### pgegp08 – Last Reached Egp Value (Erikson, Goldthorpe, Portocarero)

1	[ I] Higher Managerial and Professional Workers	25929
2	[ II] Lower Managerial and Professional Workers	39102
3	[ IIIa] Routine Clerical Work	24711
4	[ IIIb] Routine Service and Sales Work	19770
5	[ IVa] Small Self-Employed With Employees	3010
6	[ IVb] Small Self-Employed Without Employees	3462
7	[ V] Manual Supervisors	373
8	[ VI] Skilled Manual Workers	16626
9	[ VIIa] Semi- and Unskilled Manual Workers	30135
10	[ VIIb] Agricultural Labour	1752
11	[ IVc] Self-Employed Farmers	603
-1	No answer	0
-2	Does not apply	98217
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0

-7	Only available in less restricted edition	0
-8	Question this year not part of survey	507520

Waves: 2013–

This variable gives the (last reached) occupational class for all employed persons. PGEGP08 is based on ISCO-08 with is converted to ISCO-88 in a first step. Both crosswalks use the Stata ado derivescores by Daniel Bela and Knut Wenzig and are based on work of Ganzeboom (2010) and Ganzeboom, De Graaf, Treiman (1992).

Please also see occupational status (ISEI) and occupational prestige scores (SIOPS, MPS).

Harry B.G. Ganzeboom, Paul M. De Graaf, Donald J. Treiman (1992), *A standard international socio-economic index of occupational status*, in: *Social Science Research* 21 (1), 1-56, [https://doi.org/10.1016/0049-089X\(92\)90017-B](https://doi.org/10.1016/0049-089X(92)90017-B). – Ganzeboom, Harry B.G.; Treiman, Donald J., “*International Stratification and Mobility File: Conversion Tools*.” Amsterdam: Department of Social Research Methodology, <http://www.harryganzeboom.nl/ismf/index.htm>. <2010-01-12>.

For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, [kwenzig@diw.de](mailto:kwenzig@diw.de))

### **pgisei88** – Last Reached Isei Value (International Socio-Economic Index)

---

16		17842
19		2300
20		9815
21		2275
22		570
23		11897
24		1589
25		14752
26		5969
27		3337
28		2476
29		22756
30		31530
31		3182
32		4981
...	(39 rows omitted)	332215
79		368
82		1030
83		202
85		2310
87		838
88		3343
90		296
-1	No answer	1453
-2	Does not apply	180637
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	113247

Waves: 1984-2017

This variable reflects the Standard International Socio-Economic Index of Occupational Status for all employed persons. The ISEI Index was developed in 1992 by Ganzeboom, De Graaf, Treiman, and De Leew based on information about income, education, and occupation. Technically, ISEI was created by scaling the ISCO-88 classification. The values for the variable range between 16 and 90. In contrast to the prestige scores of Ganzeboom and Treiman (1996) and Wegener (1988), ISEI is a measure of socio-economic status.

It is derived from the ISCO-88 code of the current occupation using the Stata ado `iskoisei` by John Hendrickx which itself is based on Harry Ganzeboom's SPSS algorithms.

Also available: occupational prestige scores (SIOPS, MPS) and occupational class (EGP).

For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, [kwenzig@diw.de](mailto:kwenzig@diw.de))

### pgmps92 - Last Reached Mps Value (Magnitude-Prestige-Skale, Wegener)

30	205
30.1000003814697	901
30.2000007629395	2136
30.2999992370605	249
31	5424
31.1000003814697	1287
31.2000007629395	687
31.5	6298
31.7000007629395	4423
31.7999992370605	546
31.8999996185303	264
32	549
32.0999984741211	5150
32.2000007629395	645
32.2999992370605	14477
... (163 rows omitted)	422555
139.800003051758	394
145.699996948242	2220
152.5	2365
153.5	162
191.300003051758	3343
207.199996948242	614
216	781
-1 No answer	1536
-2 Does not apply	180752
-3 Not valid	0
-4 Forbidden Multiple Response	0
-5 Not Included In Questionnaire Version	0
-6 Questionnaire Version With Modified Filter	0
-7 Only available in less restricted edition	0
-8 Question not part of the survey program this year	113247

Waves: 1984-2017

This variable gives the occupational prestige score developed by Wegener (1988) for all employed persons. Like the SIOPS prestige score, Wegener's prestige scale measures a person's occupational prestige and was developed especially for use in Germany. MPS is assigned based on the German Federal Statistical Office's occupational classification of 1992 (KldB 92). The procedure has been documented in Frietsch and Wirth (2001).

Also available: occupational prestige scores (SIOPS, ISEI) and occupational class (EGP).  
 For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

### pgsiops88 – last reached SIOPS score (stand. internat. occ. prestige scale, Treiman)

---

6	12
13	435
15	2208
17	227
18	90
19	10286
20	7488
21	17491
22	3975
23	4480
24	1637
25	6242
26	1951
27	876
28	4947
... (40 rows omitted)	396032
70	6807
71	1030
72	1829
73	1529
75	194
76	296
78	5811
-1 No answer	1453
-2 Does not apply	180637
-3 Not valid	0
-4 Forbidden Multiple Response	0
-5 Not Included In Questionnaire Version	0
-6 Questionnaire Version With Modified Filter	0
-7 Only available in less restricted edition	0
-8 Question not part of the survey program this year	113247

Waves: 1984-2017

This variable gives the occupational prestige score index for all employed persons. It is based on ISCO-88 and was developed by Donald Treiman et al. The scale ranges from 6 to 78. The algorithm is based on Fritsche and Wirth (2001).

Please also see occupational prestige scores (MPS), occupational status (ISEI), and occupational class (EGP).

*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)

### pggegp88 – Last Reached Egp Value (Erikson, Goldthorpe, Portocarero)

---

1	[ I] Higher Managerial and Professional Workers	56923
2	[ II] Lower Managerial and Professional Workers	102206
3	[ IIIa] Routine Clerical Work	59349
4	[ IIIb] Routine Service and Sales Work	61432
5	[ IVa] Small Self-Employed With Employees	6845
6	[ IVb] Small Self-Employed Without Employees	13487
7	[ V] Manual Supervisors	0
8	[ VI] Skilled Manual Workers	88767
9	[ VIIa] Semi- and Unskilled Manual Workers	77630
10	[ VIIb] Agricultural Labour	6540
11	[ IVc] Self-Employed Farmers	2694
-1	No answer	1453
-2	Does not apply	180637
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	113247

Waves: 1984-2017

This variable gives the occupational class for all employed persons. It is derived from the ISCO-88 classification using Hendrickx's (2002) Stata ado. In addition, it is based on information on self-employment and number of employees (supervisory status). The EGP Index was documented by Ganzeboom/Treiman in 1996 and revised in 2003.

Information about supervisory status (number of employees of self-employed persons) is available from wave Q (2000) on. (This could lead to some minor longitudinal inconsistencies.)

Based on the new classification developed by Ganzeboom/Treiman (2003), several ISCO values were recoded in this variable 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 (PGISEI88) and occupational prestige scores (PGSIOPS88, PGMP92).

John Hendrickx, 2002. "ISKO: Stata module to recode 4 digit ISCO-88 occupational codes," *Statistical Software Components S425802*, Boston College Department of Economics, revised 20 Oct 2004. <https://ideas.repec.org/c/boc/bocode/s425802.html>

Ganzeboom, Harry B.G.; Treiman, Donald J., "International Stratification and Mobility File: Conversion Tools." Amsterdam: Department of Social Research Methodology, <http://www.harryganzeboom.nl/isco88/>.

For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, [kwenzig@diw.de](mailto:kwenzig@diw.de))

### pgerljob – Working In Occupation Trained For

1	Yes	231821
---	-----	--------

2	No	147473
3	In Training	24879
4	Has No Job Training	28445
-1	No answer	18120
-2	Does not apply	320472
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

*Waves:* all

This variable is designed to offer annual data on all employed individuals, indicating whether they are working in the occupation they were trained for. The main variable for the generation process is therefore p1b0072 (Training Applies To Current Occupation) from the pl dataset. PGERLJOB corrects the original variable by adding information from the youth and migration sample and enriching the data with information from the previous year in case of missing information if there was no occupational change.

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

### pgausb – Required Training For Job

---

1	No Training	29759
2	Intro. To Job	51286
3	On-The-Job Training	29172
4	Courses	15205
5	Vocational Training	209712
6	Technical School, Engineering (East) 90-96	2002
7	Technical College or University, until 98	10446
8	Technical College since 99	30644
9	University since 99	45719
-1	No answer	20688
-2	Does not apply	326570
-4		7
-5	Not Included In Questionnaire Version	0
-7	Only available in less restricted edition	0

*Waves:* all

This variable is designed to provide annual data on required job training for all employees. It is generated mainly from p1b0073\_v# variables, which ask about required formal education. However, it should be noted that this question is asked only every other year to all employed respondents, while people who change jobs or first-time respondents get posed this question every year. This means that not all respondents have to answer this particular question on an annual basis. Up to 2018, three biannually asked questions on required on-the-job training (variables p1b0074\_v1-v3) with 0/1 coding were additionally used. However, these questions were discontinued afterwards.

For better consistency across years, information from the previous year is carried forward if no job change occurred.

The answer option “completed technical school” was only used from 1990 to 1993 in the East German version of the questionnaire. As noted above, not all employed people are asked

annually about their required job training, therefore the value (6) is valid until 1996 for some respondents due to forward imputation.

The code (-2) is assigned to all non-employed respondents and also includes people in occupational training, in occupational retraining programs, and those doing an internship at the time of the survey.

For more information, contact: Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

### pgerwzeit – Length Of Time With Firm

-1	No answer	3777
-2	Does not apply	319821
-3	Not valid	102
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

Waves: all

The variable is designed to offer data on the duration of employment at the firm at the time of the interview for all employees. It is generated from the two variables p1b0036\_h and p1b0035, indicating respectively the respondent's starting year and month with their 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.

It should be noted that in the final variable the monthly and annual information is rounded as duration in years. This means that the duration is presented as a decimal, with months expressed as a fraction of a whole year. For example, if someone has worked for 2 years and 6 months in a firm, this is expressed as 2.5 years.

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

### pgtatzeit – Actual Work Time Per Week

---

-1	No answer	22293
-2	Does not apply	317040
-3	Not valid	1135
-7	Only available in less restricted edition	0

*Waves:* all

This variable provides data on actual weekly working hours (including overtime) for all employees at the time of the survey (including the self-employed). The data is obtained by asking respondents how many hours they work on average per week, represented in the plb0186\_v# variables from the pl dataset.

For implausible answers (actual weekly working hours of more than 80 per week), we assign the value (-3). Furthermore, the final variable is rounded as duration in hours. This means that the duration is presented as a decimal, with minutes expressed as a fraction of an hour. For example, if a respondent's weekly working time is 38 hours and 30 minutes, it is expressed as 38.5 hours.

Please also see PGVEBZEIT and PGUEBSTD.

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

### pgvebzeit – Agreed Upon Work Time Per Week

---

-1	No answer	47081
-2	Does not apply	356502
-3	Not valid	49
-7	Only available in less restricted edition	0

*Waves:* all

This variable provides data on agreed weekly working hours for those individuals who were in dependent employment (excluding self-employed) at the time of the survey. It is mainly generated from the annual question about a respondent's agreed-upon weekly working time reported in the plb0176\_v# variables from the pl dataset.



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. If individuals 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 as duration in hours. This means that the agreed hours are presented as a decimal, with minutes expressed as a fraction of an hour.

Please also see PGTATZEIT and PGUEBSTD.

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

### pguebstd – Overtime Per Week

---

0	No Overtime Last Month	209279
-1	No answer	16503
-2	Does not apply	367802
-3	Not valid	150
-7	Only available in less restricted edition	0

*Waves:* all

This variable provides data on overtime per week for all respondents 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, which is reported in the plb0193 variable in the pl dataset. The number of monthly overtime hours is then converted into weekly overtime hours by dividing the stated number of hours by 4.3. Since PGUEBSTD refers to weekly overtime during the last month, the number may deviate from the difference between average actual weekly working hours and the agreed weekly working hours.

In the years 1984, 1985 and 1987, respondents were not asked about number of hours of overtime per week. For these years PGUEBSTD was 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. If individuals 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 as duration in hours. This means that the overtime hours are presented as a decimal, with minutes expressed as a fraction of an hour.

Please also see PGVEBZEIT and PGTATZEIT.

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

### pgoeffd – Civil Service

---

1	Yes	104631
2	No	318910
-1	No answer	30625
-2	Does not apply	317044

-7 Only available in less restricted edition 0

*Waves:* all

This variable is a status variable and is designed to provide yes-or-no information for all employed individuals on whether they are employed in the public service. It is mainly generated from the p1b0040 variable, which asks about whether the respondent works for a public sector employer. If no job change occurred, information from the previous year is carried forward.

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

### **pgnace** - Industry Occupation [ pbra] (NACE Rev. 1.1, Sector)

1	Crop and animal production, hunting	6000
2	Forestry, Logging, Related Service activities	550
5	Fishing, Operation Of Fish Hatcheries And Fish Farms	33
10	Mining Of Coal And Lignite; Extraction Of Peat	1057
11	Extraction Of Crude Petroleum And Natural Gas	110
12	Mining Of Uranium And Thorium Ores	0
13	Mining Of Metal Ores	14
14	Other Mining And Quarrying	215
15	Manuf Food Products And Beverages	7506
16	Manuf Tobacco Products	87
17	Manuf Textiles	2614
18	Manuf Wearing Apparel; Dressing And Dyeing Of Fur	937
19	Tanning,Dressing Of Leather; Manuf luggage, Footwear	435
20	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	1508
21	Manuf Pulp, Paper And Paper Products	1439
...	(40 rows omitted)	309699
91	Activities of membership organizations n.e.c.	4244
92	Recreational, Cultural And Sporting Activities	5373
93	Other Service Activities	3462
95	Private Households With Employed Persons	1524
96	Undifferentiated goods producing activities of private households for own use	0
97	Undifferentiated services producing activities of private households for own use	0
99	Extra-territorial Organizations And Bodies	262
-1	No answer	2823
-2	Does not apply	308071
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	113247

*Waves:* 1984-2017

This variable is designed to provide annual data on the industry of economic activity for the company of 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. If they don't report a jobchange, information from the previous year is used, cf. introduction to section in PDF-documentation for this file.

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).

*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](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](mailto:kwenzig@diw.de))

### **pgnace2** – Industry Occupation [ pbra] (NACE Rev. 2, Sector)

1	Crop and animal production, hunting and related service activities	1517
2	Forestry and logging	204
3	Fishing and aquaculture	15
5	Mining of coal and lignite	73
6	Extraction Of Crude Petroleum And Natural Gas	55
7	Mining Of Metal Ores	0
8	Other Mining And Quarrying	46
9	Mining support service activities	1
10	Manufacture of food products	2801
11	Manufacture of beverages	208
12	Manuf Tobacco Products	14
13	Manuf Textiles	459
14	Manuf Wearing Apparel; Dressing And Dyeing Of Fur	219
15	Manufacture of leather and related products	66
16	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	498
...	(66 rows omitted)	131504
93	Sports activities and amusement and recreation activities	615
94	Activities of membership organisations	1722
95	Repair of computers and personal and household goods	160
96	Other personal service activities	1373
97	Private Households With Employed Persons	697
98	Undifferentiated goods- and services-producing activities of private households for own use	0
99	Extra-territorial Organizations And Bodies	58
-1	No answer	1645
-2	Does not apply	119740
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question this year not part of survey	507520

Waves: 2013–

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. If they don't report a jobchange, information from the previous year is used, cf. introduction to section in PDF-documentation for this file. 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. (Available in this classification since 2013 – wave bd.)

For more information, contact: Knut Wenzig (Tel. +49 30 89789 341, kwenzig@diw.de)

### pgbetr – Size of the Company

---

1	Lt 5	37867
2	Ge 5 Lt 10	22009
3	Ge 11 Lt 20	19091
4	Until 90: Lt 20	9940
5	91-04: Ge 5 Lt 20	25518
6	Ge 20 Lt 100	57489
7	Ge 100 Lt 200	26976
8	Until 98: Ge 20 Lt 200	30259
9	Ge 200 Lt 2000	87439
10	Ge 2000	99875
11	Self-Employed Without Coworkers	17180
-1	No answer	20839
-2	Does not apply	316458
-5		270
-7	Only available in less restricted edition	0

Waves: all

This variable provides data on the total size of the company for all working responders (including self-employed and employees). It is mainly generated from the annual question about the number of people employed in the entire company referring to all company sites not just the local branch. This information is reported in the p1b0049\_v# variables from the pl dataset.

Special attention should be given to categories (4), (5), and (8), which became necessary due to changes in the answer options over the years. Subsequently, the question asked about company size became more detailed. In 1991, the introduction of the category “under 5 employees” led to the creation of category (4) for the period before 1990. The wording then remained unchanged until 1999 when the option “20-100 employees” was added, resulting in the creation of category (8) for the period up to 1998. Starting from 2005, the prior category “5 to 20 employees” was further divided into two more specific categories, necessitating the creation of category (5) for the period between the initial reformulation in 1990 and 2004.

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. The enriched questionnaire revealed that in previous interviews, respondents 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 individuals were identified, and their 2012 original value of the entire company size was

replaced by their value of the local establishment size. Furthermore, data of people 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 PGALLBET for a broader categorization of the firm size, which is appropriate for analyses that include all sample years.

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

### pgallbet – Core Category Size Of The Company

---

1	LT 20	114425
2	GE 20 LT 200	114724
3	GE 200 LT 2000	87439
4	GE 2000	99875
5	SelfEmployed Without Coworkers	17180
-1	No answer	20839
-2	Does not apply	316458
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	270
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

*Waves:* all

This variable is designed to provide annual data on the core size category of the company for all employed respondents. It is generated from the input variables p1b0049\_v# regarding the total number of people employed in the entire company. Since respondents were asked about company size in increasingly detailed form over the years (see also PGBETR), the variable PGALLBET contains the lowest common denominator of the variable PGBETR, i.e., the firm size categories available across all SOEP waves. This broader categorization corresponds to the PGBETR values between 1984 and 1990 and offers a variable that is consistent across all waves.

In order to maintain longitudinal consistency, modifications were made after a one-time change in the questionnaire in 2012. Please also see PGBETR for more information.

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

## 8 Last Occupation

### pgjobend – Reasons for occupational change

---

1	Terminated by employer	11861
2	Limited Term Employment Relationship (1985-1998)	1823

3	Education, training completed (1985-1998)	763
4	Own resignation	18268
5	Mutual termination (1985-1990, since 1999)	4371
6	Employee requested transfer (1985-1998)	449
7	Company transferred employee (1985-1998)	494
8	Ended self-employment	2113
9	Temporary contract or education/training completed (since 1999)	7599
10	Took early retirement (1987-1998)	545
11	Company closed down (1991-1998, since 2001)	3775
12	Old-age pension (since 1991)	3899
13	Leave of absence/sabbatical (1999-2010)	1928
14	Leave, maternity leave and parental leave (1991-1998, since 2011)	3469
15	Other incl. early retirement, company closed, old-age pension, leave of absence/sabbatical (1985-1986)	173
16	Other incl. company closed, old-age pension, leave of absence/sabbatical (1987-1990)	130
17	Other incl. mutual termination (1991-1998)	688
18	Work permit not extended (since 2019)	26
-1	No answer	7192
-2	Does not apply	651882
-3	Not valid	8129
-4	Forbidden Multiple Response	2
-5	Not Included In Questionnaire Version	24362
-6	Questionnaire Version With Modified Filter	5024
-7	Only available in less restricted edition	0
-8		12245

#### Waves: 1985–

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

The variable is generated mainly from p1b0304\_v# variables, which ask about the reasons of the job termination. PGJOBEND offers data from interview date to interview date not from one year to the following. However, in the questionnaire respondents are asked about their occupational change since the beginning of the last year. This means that there is a possibility of respondents referring to the same occupational change in two consecutive interviews. On the basis of the month in which the job termination occurred and the interview month, duplicate answers are considered only once and the older statement is dominant. If a respondent stated a job termination in the current interview that occurred before the interview date in the previous year but didn't report it in the previous interview, the termination counts 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 a respondent states more than one reason for job termination, the first of the stated reasons is dominant.

The year 1990 is special due to the introduction of the East German sample. Since the East German questionnaire in 1990 did not contain information on the reasons for end of the job, all the observations in Sample C in for that year were assigned the value (-5).

Special attention should be given to 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, but from years 1985–1998 the

answer “Other” was additionally possible. While all explicit reasons have been recoded to uniform values, the answer “Other” in some years includes reasons for which in other years was separately asked for. Hence, “Other” was coded (15) for years 1985 and 1986, (16) for years 1987–1990, and (17) for years 1991–1998.

Note that for years 1991–1998 and again starting from 2011 there is a variable value (14) “Leave, maternity leave and parental leave” whereas for 1999–2010 the given reason was only covered in value (13) “Leave of absence/sabbatical”. In addition, codes (2) “Limited Term Employment Relationship” and (3) “Education, training completed” for years 1985–1998 were combined into to code (9) “Temporary contract or education/training completed” since 1999.

Respondents of the migration samples (M1-M6) are not asked about an occupational change in their initial interview, but in their follow-up interview the following year. Therefore, information on PGJOBEND for the first year of a migration sample are taken from the subsequent year’s interview.

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 people represented by the youth questionnaires the negative value (-5) has been introduced starting from year 2006.

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

## 9 Employment History

### pgexpft – Working Experience Full-Time Employment

-1	No answer	12562
-2	Does not apply	0
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

*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). PGEXPFT is generated by adding up the years of full-time employment before a respondent’s first interview (from PBIOSPE) plus months of full-time employment after their first interview (from ARTKALEN). 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. For annual data, if the year in which a spell started and ended is the same, a spell of 0.5 years is assumed.

PGEXPFT gives the length of full-time employment in years with months in decimal form. This means that a person with 2 years and 6 months of full-time experience, has a PGEXPFT value of 2.5 years.

Individuals 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. Respondents 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 people whose life course has not been observed completely. People with inconsistent information receive a (-3).

Starting from 2021 a new spell-type was introduced in the ARTKALEN dataset, indicating when a respondent was in short-time employment (Kurzarbeit). Since for these spells there is no information on what kind of working-contract the respondent had, but this time formally still counts as work-experience, past information was used. Respondents short-time spells are counted as full-time experience when the last available employment spell before the short-time spell was in full-time. Short-time spells were already once introduced in the years 1991 to 1995. Back then, all short-time spells were counted as full-time experience. For consistency reasons this was kept this way for these years in all waves.

Please also see PGEXPPT and PGEXPUE.

For more information, contact: Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

### pgexppt – Working Experience Part-Time Employment

0		435574
0.0829999968409538		5539
0.166999995708466		5991
0.25		5814
0.333000004291534		4886
0.416999995708466		4286
0.5		15899
0.583000004291534		4182
0.666999995708466		3973
0.75		3487
0.833000004291534		3089
0.916999995708466		3155
1		23361
1.08299994468689		3130
1.16700005531311		3454
...	(562 rows omitted)	232794
50		22
50.0830001831055		1
50.3330001831055		1
50.4169998168945		1
50.75		1
51		4
51.3330001831055		2
51.5830001831055		2
-1	No answer	12562
-2	Does not apply	0
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

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). PGEXPPT is generated by adding up the years of part-time employment before a respondent's first interview (from PBIOSPE) plus months of part-time employment after their first interview (from ARTKALEN). 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. For annual data, if the year in which a spell started and ended is the same, a spell of 0.5 years is assumed. PGEXPPT gives the length of part-time employment in years with months in decimal form.

Individuals 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. Respondents whose life course has been observed completely but with no spell of part-time employment are assigned the code (0). The code (-1) is assigned to all individuals whose life course has not been observed completely. People with inconsistent information receive a (-3).

Starting from 2021 a new spell-type was introduced in the ARTKALEN dataset, indicating when a respondent was in short-time employment (Kurzarbeit). Since for these spells there is no information on what kind of working-contract the respondent had, but this time formally still counts as work-experience, past information was used. Respondents short-time spells are counted as part-time experience when the last available employment spell before the short-time spell was in part-time.

Please also see PGEXPFT and PGEXPUE.

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

### pgexpue - Unemployment Experience

---

0	502672
0.0829999968409538	11089
0.166999995708466	9456
0.25	8325
0.333000004291534	6667
0.416999995708466	6042
0.5	33418
0.583000004291534	5839
0.666999995708466	5919
0.75	5437
0.833000004291534	4559
0.916999995708466	4130
1	23971
1.08299994468689	4529
1.16700005531311	4125
... (364 rows omitted)	122426
37	6
37.1669998168945	1
38	29
38.0830001831055	1

39		2
41		3
44		1
49		1
-1	No answer	12562
-2	Does not apply	0
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

*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). PGEXPUE is generated by adding up the years of unemployment before a respondent's first interview (from PBIOSPE) plus the months of unemployment after their first interview (from ARTKALEN). 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. For annual data, if the year in which a spell started and ended is the same, a spell of 0.5 years is assumed. PGEXPUE gives the length of time in unemployment in years with months in decimal form.

Individuals 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. Individuals 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 people whose life course has not been observed completely. People with inconsistent information receive a (-3).

Please also see PGEXPFT and PGEXPPT.

*For more information, contact:* Mattis Beckmannshagen (Tel. +49-30-89789-321, mbeckmannshagen@diw.de)

## 10 Volunteering

### pghonor1 – First Voluntary Work

---

1	Yes	15179
2	No	7449
-1	No answer	191
-2	Does not apply	100194
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	648197
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	0

Waves: 2017–

The variable PGHONOR1 gives information whether a first side job is a honorary post or not. This variable is a direct copy of the original variable in \$P.

For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

### **pghonor2** – Second Voluntary Work

---

1	Yes	5158
2	No	1002
-1	No answer	144
-2	Does not apply	116709
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	648197
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	0

Waves: 2017–

The variable PGHONOR2 gives information whether a first side job is a honorary post or not. This variable is a direct copy of the original variable in \$P.

For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

### **pghonor3** – Third Voluntary Work

---

1	Yes	1640
2	No	245
-1	No answer	63
-2	Does not apply	121065
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	648197
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	0

Waves: 2017–

The variable PGHONOR3 gives information whether a first side job is a honorary post or not. This variable is a direct copy of the original variable in \$P.

For more information, contact: Markus Grabka (Tel. +49-30-89789-339 / mgrabka@diw.de)

## **11 School, Higher and Vocational Education**

### **pgiscd97** – ISCED-1997-Classification

---

0	in school	19630
1	inadequately	27962
2	general elementary	110392
3	middle vocational	360257

4	vocational + Abi	43845
5	higher vocational	46989
6	higher education	146888
-1	No answer	15247
-7	Only available in less restricted edition	0

*Waves:* all

The educational variable (PGISCED97) 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.

*For more information, contact:* Charlotte Bartels (Tel. +49-30-89789-346)

### pgisced11 – ISCED-2011-Classification

---

0	in school	2019
1	Primary education	25709
2	Lower secondary education	41969
3	Upper secondary education	150526
4	Post-secondary non-tertiary education	27238
5	Short-cycle tertiary education	13714
6	Bachelors or equivalent level	56227
7	Masters or equivalent level	24318
8	Doctoral or equivalent level	3127
-1	No answer	3751
-2	Does not apply	422612
-7	Only available in less restricted edition	0

*Waves:* 2010–

The educational variable (PGISCED11) 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. Note that for respondents from the newly introduced migration samples, we can differentiate between university degrees (Bachelor/Master/PhD) obtained in Germany. If the university degree was obtained abroad, the degree is counted as Bachelor, because further differentiation between university degrees abroad is only possible for a subgroup of migrants.

*For more information, contact:* Charlotte Bartels (Tel. +49-30-89789-346)

**pgcasmin** – CASMIN Classification

0	(0) In School	19511
1	(1a) Inadequately Completed	29971
2	(1b) General Elementary School	87745
3	(1c) Basic Vocational Qualification	192952
4	(2b) Intermediate General Qualification	24243
5	(2a) Intermediate Vocational	159290
6	(2c_gen) General Maturity Certificate	35203
7	(2c_voc) Vocational Maturity Certificate	55648
8	(3a) Lower Tertiary Education	54694
9	(3b) Higher Tertiary Education	92194
-1	No answer	19759
-7	Only available in less restricted edition	0

*Waves:* all

Another internationally comparable educational variable is PGCASMIN where educational degrees/diplomas are classified according to the scheme “Comparative Analysis of Social Mobility in Industrial Nations (CASMIN)”. As for PGISCED, the variable is generated for all respondents retroactively from 1984 onwards and indicates the highest degree obtained by the respondent.

*For more information, contact:* Charlotte Bartels (Tel. +49-30-89789-346)

**pgbilzeit** – Amount Of Education Or Training In Years

7		26140
8.5		2843
9		98593
10		24452
10.5		167829
11		34795
11.5		108500
12		73776
13		38759
13.5		8429
14		12070
14.5		16292
15		26753
16		23184
17		2114
18		64940
-1	No answer	15301
-2	Does not apply	26440
-7	Only available in less restricted edition	0

*Waves:* all

The following statements describe the standard computation for schooling. The school years and the additional vocational training are added together.

As can be seen, the assignment table are 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 assignment is the most commonly used one when earnings functions based on human capital theory are estimated.

schooling	years
no degree	7
lower school degree	9
intermediary school	10
degree for a professional college	12
high school degree	13
other	10

additional occupational training (includes universities)	years
apprenticeship	1.5
technical schools (incl. health)	2
civil servants apprenticeship	1.5
higher technical college	3
university degree	5

For more information, contact: Peter Krause (Tel. +49-30-89789-690)

### pgpsbil – School-Leaving Degree

1	Secondary School Degree	244592
2	Intermediate School Degree	203718
3	Technical School Degree	35033
4	Upper Secondary Degree	137827
5	Other Degree	78770
6	Dropout, No School Degree	24102
7	No School Degree Yet	20187
8	No School Attended	4973
-1	No answer	12250
-2	Does not apply	9758
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	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 PGPSBILA) as well as C (see PGPSBILO) 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.

For more information, contact: Peter Krause (Tel. +49-30-89789-690)

**pgpbbil01** – Vocational Degree Received

1	Apprenticeship	312079
2	Vocational School	57624
3	Health Care School	9310
4	Technical School	41976
5	Civil Service Training	17773
6	Other degree	29096
7	Completed Vocational Training/Education in Germany	0
-1	No answer	2886
-2	Does not apply	300466
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	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. PGPBBILA) as well as C (cf. PGPBBILO) 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.

*For more information, contact:* Peter Krause (Tel. +49-30-89789-690)

**pgpbbil02** – College Degree

1	Technical College	42541
2	University, Technical College	63535
3	College Not In Germany	15111
4	Engineering, Technical School (East)	13233
5	University (East)	7973
6	graduation, state doctorate	2457
7	graduation, state doctorate (foreign country, east)	654
8	institution of higher education (youth)	0
9	Dual Studies, University of Cooperative Education	1331
10	Other Colleges (since 2014)	53
-1	No answer	8728
-2	Does not apply	615594
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	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/diplomas 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.

*For more information, contact:* Peter Krause (Tel. +49-30-89789-690)

### pgpbbil03 – No Vocational Degree

---

1	No Vocation Degree	152243
2	Apprenticeship	29282
3	University	23931
-1	No answer	10041
-2	Does not apply	555713
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

*Waves:* all

In connection with the question about vocational degrees (PGPBBIL01 and PGPBBIL02), 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.

*For more information, contact:* Peter Krause (Tel. +49-30-89789-690)

### pgpsbilo – School-Leaving Degree East Germany

---

1	8th Grade Completed	40733
2	10th Grade Completed	70621
3	College Entrance Exam	24205
4	Other Degree	966
5	Dropout, No School Degree	1970
6	No School Degree Yet	0
-1	No answer	277
-2	Does not apply	283837
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	284371
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question this year not part of survey	64230

*Waves:* all



As a supplement to the variable PGPSBIL 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.

*For more information, contact:* Peter Krause (Tel. +49-30-89789-690)

### pgpbbilo – Vocational Degree Received East Germany

---

1	Vocational Training	73306
2	Master Craftsman	7850
3	Engineering, Technical Degree	17844
4	Other degree	1176
-1	No answer	0
-2	Does not apply	529208
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	77596
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question this year not part of survey	64230

*Waves:* all

To supplement the variable PGPBBIL01 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 PGPBBIL03 if this degree has not been replaced by a more recently attained, higher-level university or college degree).

*For more information, contact:* Peter Krause (Tel. +49-30-89789-690)

### pgpsbila – School-Leaving Degree Outside Germany

---

1	School, No Degree	23476
2	School, With Degree	45954
3	Vocational Extension School	37609
4	School Leaving Degree[ Sbil] Acquired Abroad	30
5	Certificate from a different school	721
-1	No answer	973
-2	Does not apply	653208
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	9239
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

*Waves:* all

As a supplement to the PGPSBIL, this variable provides annually updated data on the highest secondary school degree/diploma attained abroad.

For more information, contact: Peter Krause (Tel. +49-30-89789-690)

### pgpbvila - Vocational Degree Outside Germany

---

1	On-The-Job Training	3272
2	Vocational Training	4472
3	Vocational School	3996
4	College	3331
5	Other	946
6	Vocational Degree[ Bbil01] Acquired Abroad	0
7	College Education[ Bbil02] Acquired Abroad	0
8	Completed Vocational Training/Education Other Country	1999
9	graduation, state doctorate (foreign country)	45
11	With Certificate, On-The-Job Training	1363
12	With Certificate, Vocational Training	6669
13	With Certificate, Vocational School	9123
14	With Certificate, University	16238
15	With Certificate, Other	1383
16	With Certificate, Vocational Degree[ bbil01] Acquired Abroad	131
17	With Certificate, College Education[ Bbil02] Acquired Abroad	316
18	With Certificate, Completed Vocational Training/Education Other Country	0
19	With Certificate, PhD (Other Country)	511
-1	No answer	1239
-2	Does not apply	707957
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	8219
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

Waves: all

As a supplement to the variable PGPBBIL01, this variable gives (and updates) the highest-level vocational degree attained abroad.

For more information, contact: Peter Krause (Tel. +49-30-89789-690)

### pgfield - Field of tertiary education

---

1	Linguistics and cultural studies	562
2	Protestant theology	7
3	Catholic theology	36
4	Philosophy, ethics	342
5	History	535
6	Library science, documentation, journalism	265
7	General and comparative literature	904
8	Classical philology, modern greek	164
9	German philology	1801
10	English studies	716

11	Roman philology	295
12	Slavic, Baltic, Finno-Ugrian Studies	133
13	Non-european linguistics and cultural studies	181
14	Cultural studies	118
15		29
...	(87 rows omitted)	69380
303	Communication science	1
304	Media management	1
310	Renewable energies	1
320	Nutritional science	1
365	Early childhood education	2
370	Economic engineering with engineering focus	3
464	Facility Management	1
996	Undefined	6
-1	No answer	71335
-2	Does not apply	624025
-3	Not valid	366
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	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 PGPBBIL02. While the latter variable records if a person holds a degree PGFIELD contains more detailed information on the type of the degree. The data of the generated variable PGFIELD stem from two sources:

1. Person questionnaire: Each year since 1985 respondents are asked if they have left education since the beginning of the year prior to the survey and which degrees they have obtained. This part of the questionnaire contains an open question on the type and the field of newly obtained tertiary degrees. This information is coded and used for the generation of the variable PGFIELD.
2. Biography questionnaire: Since 2001 similar information is collected from respondents who fill in the biography questionnaire (usually during the first two years of participation in the panel). In contrast to the information from the person questionnaire the questions do not refer to currently obtained degrees but to degrees obtained during the time before being part of the SOEP sample.

In the variable PGFIELD we combine these two types of information. 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 PGFDT\_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 variable PGFIELD would be filled as follows:

year	value
1984-1986	-2 does not apply

year	value
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 PGFIELD we recoded the original values. As the newer version is more precise this could be done with hardly any loss of information. Some categories are collapsed. Category 3 is coded as 2 (no distinction between catholic and protestant theology), 14 as 13, 17 as 16, 24 as 23, 25 as 26 and 48 as 49. The original values of the data collected from the person questionnaire up to 2009 are stored in the respective variables in the dataset PL.

Please note that for respondents from the newly introduced migration sample a valid value to PGFIELD was only assigned if the tertiary degree was attained in Germany.

*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)*

### pgdegree – Type of tertiary degree

2	Magister	1945
5	Academically certified translator	1
6	PhD	3285
8	First state examination	4299
11	Diplom (university)	16399
15	Teacher training, BA, elementary school	1
21	Teacher training, elementary school	2142
23	Teacher training, secondary school	949
24	Teacher training, secondary school, level 1	0
25	Teacher training, academic secondary school	1355
27	Teacher training, vocational school	245
29	Other teacher training	6
30	Teacher training, MA	6630
31	Teacher training, BA	0
42	Teacher training, primary level	93
...	(11 rows omitted)	13831
84	Bachelor (technical college)	1482
87	Bachelor (vocational school)	24
88	Master (university)	2155
90	Master (technical college)	353

93	Bachelor (university) - teacher	1
94	Other degree - degree certificate	1
95	Other degree in Germany	2462
96	Final examination abroad	8
-1	No answer	89083
-2	Does not apply	624460
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	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 PGPBBIL02. While the latter variable records if a persons holds a degree, PGDEGREE contains more detailed information on the type of the degree. The data of the generated variable PGDEGREE stem from two sources:

1. Person questionnaire: Each year since 1985 respondents are asked if they have left education since the beginning of the year prior to the survey and which degrees they have obtained. This part of the questionnaire contains an open question on the type and the field of newly obtained tertiary degrees. This information is coded and used for the generation of the variables PGDEGREE.
2. Biography questionnaire: Since 2001 similar information is collected from respondents who fill in the biography questionnaire (usually during the first two years of participation in the panel). In contrast to the information from the person questionnaire the questions do not refer to currently obtained degrees but to degrees obtained during the time before being part of the SOEP sample.

In the variable PGDEGREE we combine these two types of information. However, since the retrospective information was not collected before 2001 the variable covers until 2000 only persons for whom we have prospectively observed the end of study. This explains why the number of valid observations is rather small in these years. Information on the data source is stored in the variable PGFDT\_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 PGDEGREE we recoded the original values from years 2009 and earlier. As the newer version is more precise this could be done with hardly any loss of information. Some categories are collapsed. Category 16 was mostly likely coded as 15 in earlier years, 34 as 35 and 43 as 44. The original values of the data collected from the person questionnaire 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 PGDEGREE was only assigned if the tertiary degree was attained in Germany.

*Statistisches 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)*

#### pgtraina - Apprenticeship - two-digit occupation KldB92

---

110	Farmers, general	1690
111	Fruit and vegetable farmers (non-horticultural)	12
112	Arable farmers (special, permanent crops)	42
113	Livestock farmers and pasture farmers	0
114	Seed, crop producers, propagators (non-horticultural)	7
115	Crop protectors	34
116	Farmers and landlords	61
118	Farmers and wine growers	11
120	Wine growers, general	39
121	Vine propagators	0
129	Other wine growers	0
130	Agricultural workers, general	409
131	Agricultural supervisors	0
132	Agricultural machinery drivers	0
133	Vineyard workers	0
...	(2263 rows omitted)	168287
9911	Specialized professionals without further specification	130
9921	Homeworkers without further specification	0
9931	Forepersons, group leaders without further specification	0
9941	Persons doing community service without further specification	0
9951	Self-employed persons without further specification	0
9961	Consultancy, planning professionals without further specification	13
9971	Other employees without further specification	10804
-1	No answer	4653
-2	Does not apply	583398
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0

-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	1620

*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 PGPBBIL01. In addition to the variable PGTRAINA, we provide the variables PGTRAINB, PGTRAINC and PGTRAIND. All these variables record the occupation of vocational training. The difference is that PGTRAINA contains information on vocational training within the German dual system which combines firm-based and school-based training (apprenticeship). PGTRAINB is designed to provide information on the occupation of full-time school based vocational training. PGTRAINC contains information on level vocational training (e.g., Meister, Techniker). PGTRAIND is designed to provide information on the occupation of civil servant training (“Beamtensausbildung”). We describe in brief detail the construction of the variables PGTRAINA, PGTRAINB, PGTRAINC and PGTRAIND are constructed in an analogous manner.

The data of the generated variable PGTRAINA stem from two sources:

1. Person questionnaire: Each year since 1985 respondents are asked if they have left education since the beginning of the year prior to the survey and which degrees they have obtained. This part of the questionnaire contains an open question on the type and the field of newly obtained tertiary degrees. This information is coded and used for the generation of the variable PGTRAINA.
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 PGTRAINA 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 PGFDT\_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 PGTRAINA 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 PGTRAINA 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 PL. Please note that for respondents from the newly introduced migration sample a valid value to PGTRAINA was only assigned if the vocational training was completed in Germany.

*Hartmann/Schütz (2002): Die Klassifikation der Berufe und der Wirtschaftszweige im Sozio-ökonomischen Panel. Neuvercodung der Daten 1984–2001. Infratest Sozialforschung, München. For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)*

### pgtrainb – Vocational school - two-digit occupation KldB92

110	Farmers, general	56
111	Fruit and vegetable farmers (non-horticultural)	0
112	Arable farmers (special, permanent crops)	0
113	Livestock farmers and pasture farmers	0
114	Seed, crop producers, propagators (non-horticultural)	0
115	Crop protectors	0
116	Farmers and landlords	8
118	Farmers and wine growers	0
120	Wine growers, general	0
121	Vine propagators	0
129	Other wine growers	0
130	Agricultural workers, general	45
131	Agricultural supervisors	0
132	Agricultural machinery drivers	0
133	Vineyard workers	0
...	(2263 rows omitted)	29398
9911	Specialized professionals without further specification	54
9921	Homeworkers without further specification	0
9931	Forepersons, group leaders without further specification	0
9941	Persons doing community service without further specification	0
9951	Self-employed persons without further specification	0
9961	Consultancy, planning professionals without further specification	0
9971	Other employees without further specification	5220
-1	No answer	1441
-2	Does not apply	733417
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	1571



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 PGTRAINA for more details on the construction and the values of the variable.

For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

### pgtrainc – Higher vocational school - two-digit occupation KldB92

---

110	Farmers, general	403
111	Fruit and vegetable farmers (non-horticultural)	0
112	Arable farmers (special, permanent crops)	42
113	Livestock farmers and pasture farmers	0
114	Seed, crop producers, propagators (non-horticultural)	0
115	Crop protectors	0
116	Farmers and landlords	0
118	Farmers and wine growers	0
120	Wine growers, general	20
121	Vine propagators	0
129	Other wine growers	0
130	Agricultural workers, general	14
131	Agricultural supervisors	0
132	Agricultural machinery drivers	0
133	Vineyard workers	0
...	(2263 rows omitted)	19009
9911	Specialized professionals without further specification	0
9921	Homeworkers without further specification	0
9931	Forepersons, group leaders without further specification	0
9941	Persons doing community service without further specification	0
9951	Self-employed persons without further specification	0
9961	Consultancy, planning professionals without further specification	0
9971	Other employees without further specification	3356
-1	No answer	988
-2	Does not apply	746231
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	1147

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 PGTRAINA for more details on the construction and the values of the variable.

For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

### pgtraind – Civil servant training - two-digit occupation KldB92

---

110	Farmers, general	2
111	Fruit and vegetable farmers (non-horticultural)	0
112	Arable farmers (special, permanent crops)	0
113	Livestock farmers and pasture farmers	0
114	Seed, crop producers, propagators (non-horticultural)	0
115	Crop protectors	0
116	Farmers and landlords	0
118	Farmers and wine growers	0
120	Wine growers, general	0
121	Vine propagators	0
129	Other wine growers	0
130	Agricultural workers, general	0
131	Agricultural supervisors	0
132	Agricultural machinery drivers	0
133	Vineyard workers	0
...	(2263 rows omitted)	6762
9911	Specialized professionals without further specification	0
9921	Homeworkers without further specification	0
9931	Forepersons, group leaders without further specification	0
9941	Persons doing community service without further specification	0
9951	Self-employed persons without further specification	0
9961	Consultancy, planning professionals without further specification	0
9971	Other employees without further specification	1988
-1	No answer	421
-2	Does not apply	761827
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	210

*Waves: -1985*

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

*For more information, contact:* Charlotte Bartels (Tel. +49-30-89789-346)

### **pgfdt\_f** – Data source FIELD, DEGREE, TRAIN

---

1	Individual Questionnaire	21625
2	Gap Questionnaire (temporary drop-outs)	824
3	Biographical Questionnaire	213301
4	Various Sources	0
-1	[ de] No Answer	0
-2	Does not apply	535460
-7	Only available in less restricted edition	0

*Waves: -1985*

This is a flag variable which provides information on the data sources used for the construction of the variables PGFIELD, PGDEGREE, PGTRAINA, PGTRAINB, PGTRAINC and PGTRAIND (see the description of the respective variables for details).

For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

### **pgbiltzch** – Change in Education since last survey / last year

---

0	Inconsistent educational information since last year	612813
1	Inconsistent educational information since last survey	7
2	Inconsistent educational information since last year	3564
-2	Does not apply	154826
-7	Only available in less restricted edition	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-by-case 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.

File BIOEDU contains new data on consistent longitudinally tested educational transitions.

For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

### **pgbiltztev** – Change in Education, total observed period

---

0	Consistent educational information	655235
1	Inconsistent educational decline	36
2	Inconsistent educational increase	66672
3	Inconsistent educational decline and increase	0
-2	Does not apply	49267
-7	Only available in less restricted edition	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 PGBILZTCH for more details on the sources of these inconsistencies.

For more information, contact: Charlotte Bartels (Tel. +49-30-89789-346)

## 12 Political Attitudes

### pgpicoriga – Party Identification in Country of Origin

---

1	Yes	241
2	No	4088
-1	No answer	136
-2	Does not apply	0
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	25405
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question this year not part of survey	741340

### pgpicorigb – Party Identified with in Country of Origin

---

		741340
	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	4224
	... (22 rows omitted)	25557
	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

### pgpicorigc – Party Family of PI in Country of Origin

---

0	Not Categorized	57
1	ECO_Ecological Parties	0
2	COM_Socialist Parties	7
3	SOC_Social Democratic Parties	6
4	LIB_Liberal Parties	3
5	CHR_Christian Democratic Parties	0
6	CON_Conservative Parties	10
7	NAT_Nationalist Parties	23
8	AGR_Agrarian Parties	0
9	ETH_Ethnic and Regional Parties	22
10	SIP_Special Issue Party	0
11	ISP_Islamic Parties	7
-1	No answer	106
-2	Does not apply	4224
-3	Not valid	0
-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	25405
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0
-8	Question this year not part of survey	741340

## 13 Information on the Interview

### pgmonth – Month Of Interview

---

1	January	45171
2	February	177534
3	March	192092
4	April	112956
5	May	69752
6	June	53020
7	July	38973
8	August	28891
9	September	21798
10	October	14964
11	November	9723
12	December	6287
-1	No answer	0
-2	Does not apply	0
-3	Not valid	49
-4	Forbidden Multiple Response	0

-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	0

Waves: all

Month of interview is generated using the answers to the individual questionnaire. Missing answers are filled in using data from the files HBRUTTO and HBRUTTOL. Interviews that took place in December and before the 20th of that month were recoded -3.

For more information, contact: Peter Krause (Tel. +49-30-89789-690)

### pgpiyear – Personal Related Survey Year

---

2009		20776
2010		259
2017		31894
2018		31402
2019		31169
2020		29979
2021		20906
2022		1162
-2	Does not apply	0
-7	Only available in less restricted edition	0
-8	Question this year not part of survey	603663

Waves: all

For more information, contact: Peter Krause (Tel. +49-30-89789-690)

### pgmode – Interview Method

---

100	With Interviewer Assistance	10877
110	Oral Interview	161438
120	Written Ques. No Interviewer	143392
130	Mixed Type	0
131	Written Ques. Interviewer	25248
132	Oral And Written	19728
133	Proxy	155
134	Third Person Present	3227
135	No Third Person Present	20991
140	CAPI - Since 1998 (O)	282927
150	CAWI Since 2014 (BE)	12920
160	CATI since 2021	663
170	CASI since 2021	331
200	Telephone Assistance	0
210	Written, By Mail	72817
220	Telephone Interview	4404
230	CAPI by phone	10968
-1	No answer	1105
-2	Does not apply	0
-3	Not valid	19

-4	Forbidden Multiple Response	0
-5	Not Included In Questionnaire Version	0
-6	Questionnaire Version With Modified Filter	0
-7	Only available in less restricted edition	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 files PBRUTTO and PBRUTTOL.

For more information, contact: Peter Krause (Tel. +49-30-89789-690)

### inputdataset – Source dataset

---

apgen	12245
bapgen	27124
bbpgen	29264
bcpgen	28520
bdpgen	31523
bepgen	28042
bfpgen	27743
bgpgen	29870
bhpgen	33265
bipgen	30997
bjpgen	30397
bkpgen	31083
blpgen	20770
bpgen	11090
cpgen	10646
... (8 rows omitted)	97883
lpgen	13768
mpgen	13511
npgen	13283
opgen	14670
ppgen	14085
qpgen	24576
rpgen	22351
spgen	23892
tpgen	22611
upgen	22019
vpgen	21105
wpgen	22665
xpgen	21232
ypgen	19945
zpgen	21035

Waves: all