

DIW Berlin / SOEP (Ed.)

**Research Report**

## SOEP-Core v34 - PGEN: Person-related status and generated variables

SOEP Survey Papers, No. 758

**Provided in Cooperation with:**

German Institute for Economic Research (DIW Berlin)

*Suggested Citation:* DIW Berlin / SOEP (Ed.) (2019) : SOEP-Core v34 - PGEN: Person-related status and generated variables, SOEP Survey Papers, No. 758, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/203352>

**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/>

758

## SOEP Survey Papers

Series D – Variable Descriptions and Coding

SOEP – The German Socio-Economic Panel at DIW Berlin

2019

# SOEP-Core v34 – PGEN: Person-Related Status and Generated Variables

SOEP Group

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

The aim of the SOEP Survey Papers Series is to thoroughly document the survey's data collection and data processing.

The SOEP Survey Papers is comprised of the following series:

Series A – Survey Instruments (Erhebungsinstrumente)

Series B – Survey Reports (Methodenberichte)

Series C – Data Documentation (Datendokumentationen)

Series D – Variable Descriptions and Coding

Series E – SOEPmonitors

Series F – SOEP Newsletters

Series G – General Issues and Teaching Materials

The SOEP Survey Papers are available at <http://www.diw.de/soepsurveyspapers>

Editors:

Dr. Jan Goebel, DIW Berlin

Prof. Dr. Stefan Liebig, DIW Berlin and Freie Universität Berlin

Dr. David Richter, DIW Berlin

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

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

Dr. Sabine Zinn, DIW Berlin

Please cite this paper as follows:

SOEP Group, 2019. SOEP-Core v34 – PGEN: Person-Related Status and Generated Variables. SOEP Survey Papers 758: Series D – Variable Descriptions and Coding. Berlin: DIW Berlin/SOEP



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

© 2019 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 v34 – PGEN: Person-Related Status and Generated Variables

SOEP Group

2019

The file pgen is part of a collection, which is released with doi:[10.5684/soep.v34](https://doi.org/10.5684/soep.v34).

## Contents

<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 Household Number (Household ID) . . . . .	5
	cid – Case-ID, Original Household Number . . . . .	5
	sample1 – Subsample Identifier . . . . .	5
	year – Survey Year . . . . .	6
<b>3</b>	<b>Own Nationality and Residential Status</b>	<b>7</b>
	pgnation – 1. Nationality . . . . .	7
	pgstatus_asyl – Status quo, Application for Asylum . . . . .	8
	pgstatus_refu – Refugee residence status . . . . .	8
<b>4</b>	<b>Familiy Status and Partnership</b>	<b>8</b>
	pgpartz – Partner Indicator . . . . .	8
	pgpartnr – Partner Person Number . . . . .	9
	pgfamstd – Marital Status In Survey Year . . . . .	9
<b>5</b>	<b>Wages and Salary</b>	<b>10</b>
	pglabgro – Current Gross Labor Income in Euro . . . . .	10
	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 . . . . .	12
	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 . . . . .	14
	pgimpsnd1 – Imputation flag, Current Gross Income Secondary Job (1) . . . . .	15
	pgimpsnd2 – Imputation flag, Current Gross Income Secondary Job (2) . . . . .	16
	pgimpsnd3 – Imputation flag, Current Gross Income Secondary Job (3) . . . . .	16
<b>6</b>	<b>Current Employment Status</b>	<b>16</b>
	pgstib – Occupational Position . . . . .	16
	pgemplst – Employment Status . . . . .	18
	pglfs – Labor Force Status . . . . .	19
	pgjobch – Occupational Change . . . . .	20
	pgsndtyp1 – 1. secondary job: Occupational Position . . . . .	21
	pgsndtyp2 – 1. secondary job: Occupational Position . . . . .	22
	pgsndtyp3 – 3. secondary job: Occupational Position . . . . .	22
<b>7</b>	<b>Current Occupation</b>	<b>23</b>
	pgisco88 – Current Occupational Classification (ISCO-88 Com) . . . . .	23
	pgisco08 – Current Occupational Classification (ISCO-08) . . . . .	24
	pgkldb92 – Current Occupational Classification (KldB92) . . . . .	25
	pgkldb2010 – Current Occupational Classification (KldB2010) . . . . .	26
	pgisei88 – Last Reached Isei Value (International Socio-Economic Index) . . . . .	26
	pgmps92 – Last Reached Mps Value (Magnitude-Prestige-Skale, Wegener) . . . . .	27

pgsiops88 – Last Reached Siops Value (Std. Internat. Occupational Prestige Scale), Treim . . . . .	28
pgegp88 – Last Reached Egp Value (Erikson, Goldthorpe, Portocarero) . . . . .	29
pgerljob – Working In Occupation Trained For . . . . .	30
pgausb – Required Training For Job . . . . .	31
pgerwzeit – Length Of Time With Firm . . . . .	31
pgtatzeit – Actual Work Time Per Week . . . . .	32
pgvebzeit – Agreed Upon Work Time Per Week . . . . .	33
pguebstd – Overtime Per Week . . . . .	33
pgoeffd – Civil Service . . . . .	33
pgnace – Industry Occupation [pbra] (NACE Rev. 1.1, Sector) . . . . .	34
pgnace2 – Industry Occupation [pbra] (NACE Rev. 2, Sector) . . . . .	35
pgbetr – Size of the Company . . . . .	36
pgallbet – Core Category Size Of The Company . . . . .	37
<b>8 Last Occupation</b>	<b>37</b>
pgjobend – Reasons for occupational change . . . . .	37
<b>9 Employment History</b>	<b>39</b>
pgexpft – Working Experience Full-Time Employment . . . . .	39
pgexppt – Working Experience Part-Time Employment . . . . .	39
pgexpue – Unemployment Experience . . . . .	40
<b>10 Volunteering</b>	<b>41</b>
pghonor1 – Third Voluntary Work . . . . .	41
pghonor2 – Second Voluntary Work . . . . .	42
pghonor3 – First Voluntary Work . . . . .	42
<b>11 School, Higher and Vocational Education</b>	<b>42</b>
pgiscd97 – ISCED-1997-Classification . . . . .	42
pgiscd11 – ISCED-2011-Classification . . . . .	43
pgcasmin – CASMIN Classification . . . . .	44
pgbilzeit – Amount Of Education Or Training In Years . . . . .	44
pgpsbil – School-Leaving Degree . . . . .	45
pgpbil01 – Vocational Degree Received . . . . .	46
pgpbil02 – College Degree . . . . .	46
pgpbil03 – No Vocational Degree . . . . .	47
pgpsbilo – School-Leaving Degree East Germany . . . . .	47
pgpbilo – Vocational Degree Received East Germany . . . . .	48
pgpsbila – School-Leaving Degree Outside Germany . . . . .	48
pgpbila – Vocational Degree Outside Germany . . . . .	49
pgfield – Field of tertiary education . . . . .	49
pgdegree – Type of tertiary degree . . . . .	51
pgtraina – Apprenticeship - two-digit occupation KldB92 . . . . .	53
pgtrainb – Vocational school - two-digit occupation KldB92 . . . . .	55
pgtrainc – Higher vocational school - two-digit occupation KldB92 . . . . .	56
pgtraind – Civil servant training - two-digit occupation KldB92 . . . . .	57
pgfdt_f – Data source FIELD, DEGREE, TRAIN . . . . .	57
pgbiltzch – Change in Education since last survey / last year . . . . .	58
pgbiltztev – Change in Education, total observed period . . . . .	59

<b>12 Political Attitudes</b>	<b>59</b>
pgpicoriga – Party Identification in Country of Origin . . . . .	59
pgpicorigb – Party Identified with in Country of Origin . . . . .	59
pgpicorigc – Party Family of PI in Country of Origin . . . . .	60
<b>13 Information on the Interview</b>	<b>60</b>
pgpmnth – Month Of Interview . . . . .	60
pgpiyear – Personal Related Survey Year . . . . .	61
pgpmode – Interview Method . . . . .	61
inputdataset – Source dataset . . . . .	62

## 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 Household Number (Household ID)

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

### cid – Case-ID, Original Household Number

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.

### sample1 – Subsample Identifier

1	[1] A German HH Head, West	14568
2	[2] B1 Turkish HH Head	1070
3	[3] B2 Yugoslav HH Head	550
4	[4] B3 Greek HH Head	353
5	[5] B4 Italian HH Head	712
6	[6] B5 Spanish HH Head	211
7	[7] C German HH Head, East	9324
8	[8] D1 Immigrant W. Germany (1984-1994), 1994/95	933
9	[9] D2 Immigrant W. Germany (1984-1994), 1994/95	849



10	[10] E1 Refreshment (PAPI)	1151
11	[11] E2 Refreshment (CAPI)	1065
12	[12] F1 Refreshment (Innovation)	18777
13	[13] F2 Refreshment, Additional Foreigner Addresses	462
14	[14] G1 High Income (PreTest)	0
15	[15] G2 High Income	5340
...	(9 rows omitted)	30504
31	[31] M4 2016 Refugee/Families (2013-2015)	0
60	[60] L2/3 FiD Screening, Low Income	6879
61	[61] L2/3 FiD Screening, Single Parent	4141
62	[62] L2/3 FiD Screening, Families with Several Children	7491
63	[63] L1 FiD Birth Cohort 2007	3009
64	[64] L1 FiD Birth Cohort 2008	3117
65	[65] L1 FiD Birth Cohort 2009	2887
66	[66] L1 FiD Birth Cohort 2010	3024
-1	[-1] No Answer	0
-2	[-2] Does not apply	14
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	541532

#### sy<sup>year</sup> – 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
...	(4 rows omitted)
2003	22611
2004	22019
2005	21105
2006	22665
2007	21232
2008	19945
2009	21035
2010	27124
2011	29264

2012	28520
2013	31523
2014	28042
2015	27743
2016	29870
2017	33265

### 3 Own Nationality and Residential Status

#### pgnation – 1. Nationality

---

1	[1] Germany	570155
2	[2] Turkey	22577
3	[3] Ex-Yugoslavia	7443
4	[4] Greece	7962
5	[5] Italy	11769
6	[6] Spain	4492
7	[7] Ex-GDR (Country Of Origin Only)	0
10	[10] Austria	1402
11	[11] France	666
12	[12] Benelux	119
13	[13] Denmark	154
14	[14] Great Britain	564
15	[15] Sweden	85
16	[16] Norway	31
17	[17] Finland	88
...	(173 rows omitted)	30355
191	[191] Gabon	0
192	[192] Comoros	0
193	[193] Katar	0
194	[194] Sahara	0
195	[195] Ingushetia	3
222	[222] Eastern Europe	0
333	[333] Other Unspecified Foreign Country	0
999	[999] ethnic minorities (e.g. Yazidi, Roma)	2
-1	[-1] No Answer	90
-2	[-2] Does not apply	6
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

*Waves:* all

This variable is designed to integrate the information on respondent's nationality for all subsamples. Since some members of Sample B (persons with Turkish, Italian, Spanish, Greek, and Yugoslavian citizenship) received the question items in their own language up to 1995, to carry out an integrated analysis with Sample B, the user must obtain this information from the corresponding \$PAUSL files and add it to the individual data. The variable PGNATION thus offers a variable on nationality for all subsamples.

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

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

---

1	[1] Entitlement to protection (asylum, refugee protection, residence permit)	4999
2	[2] Temporary Suspension of Deportation	513
3	[3] Request to Leave Germany	343
4	[4] Decision Regarding Residential Status Still Open	1948
-1	[-1] No Answer	103
-2	[-2] Does not apply	27316
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	27913
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	594828

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 arefbck in the dataset ppfadl.

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

**pgstatus\_refu** – Refugee residence status

---

1	[1] In Proceedings	2459
2	[2] Entitlement to protection (asylum, refugee protection, residence permit)	6381
3	[3] Temporary Suspension of Deportation	559
4	[4] Other	467
-1	[-1] No Answer	170
-2	[-2] Does not apply	25186
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	27913
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	594828

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 arefbck 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	[0] No partner	203138
1	[1] Spouse, registered partner	397840
2	[2] Partner	54498
3	[3] Probably spouse, registered partner	689
4	[4] Probably partner	1181
5	[5] not clear	617

-1	[-1] No Answer	0
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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:* Martin Kroh (Tel. +49-30-89789-678)

### pgpartnr – Partner Person Number

*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 (BHPPNR in 2017, PLK0001 in PL), the household matrix reported by the head of household at the beginning of the interview (BHSTELL in 2017, STELL\_V1, STELL\_V2, 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. 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:* Martin Kroh (Tel. +49-30-89789-678)

### pgfamstd – Marital Status In Survey Year

1	[1] Married	394268
2	[2] Married, But Separated	13764
3	[3] Single	159905

4	[4] Divorced	44143
5	[5] Widowed	38442
6	[6] husband/wife abroad	1466
7	[7] Registered Same-Sex Partnership, Living Together	474
8	[8] Registered Same-Sex Partnership, Living Apart	82
-1	[-1] No Answer	2483
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	2898
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	38
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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 still be possible. For those whose partner was identified within the household, marital status is counter-checked with the information given by the partner. Where contradictions can be found, indication of the person information is compiled if reasonable. If no information is available, the indication by position related to head of household is deferred. Remaining contradictions are solved using information on marriage status when a child was born as well as future reports on a given relationship. Marital status is only available for people, who were interviewed. Data on marital status provided here is generated from and hence consistent with the corresponding BIOMARSY 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:* Paul Schmelzer (Tel. +49-30-89789-526, pschmelzer@diw.de)

## 5 Wages and Salary

### **pglabgro** – Current Gross Labor Income in Euro

*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	[0] Not imputed	343468
1	[1] Imputed	40058
-1	[-1] No Answer	0
-2	[-2] Does not apply	269742
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	4695
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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

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	[0] Not imputed	355122
1	[1] Imputed	28403
-1	[-1] No Answer	0
-2	[-2] Does not apply	269743
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	4695
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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

*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 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	[0] Not imputed	27842
1	[1] Imputed	3112
-1	[-1] No Answer	0
-2	[-2] Does not apply	457328
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0

-5	[-5] Not included in this version of the questionnaire	169681
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

Waves: 1995-2016

The variable PGIMPSND designates imputations of itemnonresponse in the variable PGSND-JOB\$\$ (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](mailto:mgrabka@diw.de))

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

---

0	107
2	1
5	3
7	1
10	7
12	4
15	4
16	4
17	1
18	1
19	1
20	22
24	1
25	9
28	3
... (236 rows omitted)	1616
2300	1
2400	1
2500	5
2600	1
2800	1
3000	1
3500	1
3695	1
4000	5
4200	1
4980	1
5000	2
6000	2
-2	25106
-5	631049

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)

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

---

0	18
3	1
6	1
7	1
8	1
10	3
15	2
18	1
20	6
24	1
25	3
30	9
40	9
45	1
48	2
... (42 rows omitted)	155
500	6
550	1
600	4
700	3
724	1
750	1
800	2
860	1
900	2
1000	4
1200	1
1500	1
1700	1
-2	26672
-5	631049

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                      6

2	1
15	1
20	2
40	2
50	3
57	1
60	3
70	1
75	1
83	1
90	1
100	5
108	1
110	2
120	1
140	1
150	1
200	1
284	1
400	2
450	1
500	1
800	1
1000	1
2000	1
2816	1
-2	26870
-5	631049

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	[0] Not imputed	1743
1	[1] Imputed	65
-1	[-1] No Answer	0
-2	[-2] Does not apply	31314
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	624841
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

Waves: 2017–

The variable PGIMSND1 designates imputations of item non-response in the variable PGSND-JOB1 (current gross labor income from a side job).

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

### pgimsnd2 – Imputation flag, Current Gross Income Secondary Job (2)

---

0	[0] Not imputed	232
1	[1] Imputed	10
-1	[-1] No Answer	0
-2	[-2] Does not apply	32880
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	624841
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

Waves: 2017–

The variable PGIMSND2 designates imputations of item non-response in the variable PGSND-JOB2 (current gross labor income from a side job).

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

### pgimsnd3 – Imputation flag, Current Gross Income Secondary Job (3)

---

0	[0] Not imputed	40
1	[1] Imputed	4
-1	[-1] No Answer	0
-2	[-2] Does not apply	33078
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	624841
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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

---

0	[0] Do Not Know	0
10	[10] Not Employed	64734
11	[11] In Education	31418
12	[12] Unemployed, Not Employer	39260
13	[13] Pensioner	131177
15	[15] Military, Community Service	1979

110	[110] Apprentice	6981
120	[120] Apprentice, Trainee Industry Technology	6833
130	[130] Apprentice, Trainee Trade And Commerce	3956
140	[140] Student Trainee, Intern	1925
150	[150] Aspirant	3
210	[210] Untrained Worker	19527
220	[220] Semi-Trained Worker	41615
230	[230] Trained Worker	45477
240	[240] Foreman, Team Leader	6092
...	(22 rows omitted)	88734
530	[530] Qualified Professional	82508
540	[540] Highly Qualified Professional	45224
550	[550] White-Collar Worker With Managerial Functions	5693
610	[610] Low-Level Civil Service	898
620	[620] Middle-Level Civil Service	6693
630	[630] High-Level Civil Service	10305
640	[640] Executive Civil Service	6627
999	[999] Employment without information on employment status	0
-1	[-1] No Answer	5560
-2	[-2] Does not apply	4744
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

Waves: all

The variable represents a compilation of all relevant information on current occupational position. It is generated by combining information on “occupational group”, “unemployed (yes/no)”, “military/community service”, “in education (yes/no)”, and “pensioner”. A hierarchical scheme is used to determine which data is given precedence when a variety of divergent information exists (increasing dominance):

value	label
10	not employed
13	pensioner
11	currently in education
15	military / community service
12	registered unemployed
110-150	apprentice
410-440	self-employed
210-250	manual laborer
510-550	employee
610-640	civil service

The categories (150) and (310) to (340) were only assigned to respondents in East Germany in 1990. In PGSTIB, non-working persons are only assigned to the category (13) “pensioner” if they are recipients of retirement pension or if they are recipients of widow’s pension AND are older than 60 years. 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.

For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

### pgemplst – Employment Status

1	[1] Full-Time Employment	266009
2	[2] Regular Part-Time Employment	69002
3	[3] Vocational Training	19059
4	[4] Marginal, Irregular Part-Time Employment	28627
5	[5] Not Employed	274630
6	[6] Sheltered workshop	575
-1	[-1] No Answer	24
-2	[-2] Does not apply	35
-3	[-3] Answer improbable	2
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

*Waves:* all

This variable is generated from the annual question on current employment status, which has a central filter function in the questionnaire to separate employed people from non-employed people for further questions. It is designed to provide consistent longitudinal data on employment status across all waves.

Since the beginning of the SOEP in the year 1984, a consistent status variable has been used to differentiate among different types of employment status. The category “not employed” comprises non-working individuals, those in military/community service, those on maternity leave, and employed persons in a phased retirement scheme (Altersteilzeit) whose current actual working hours are zero. From 1998 on, the additional category “sheltered workshop” is included for disabled persons in sheltered employment.

PGEMPLST supplements the variable PGLFS, which differentiates among persons 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 the alternative status. Those teenagers who attend school, are in vocational preparational year or in a vocational training, or in vocational school were assigned value of 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).

Due to a change in the 2017 questionnaire, the category „[8] NW-but sometimes sec. job“ is no longer used as there is no possibility to retract whether an individual has a second job on irregular basis. If a person is not employed but has at least one secondary job, it is assumed that it is on regular basis (category 10).

For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

**pglfs** – Labor Force Status

1	[1] Non-Working	79580
2	[2] NW-Age 65 And Older	97521
3	[3] NW-In Education-Training	28022
4	[4] NW-Maternity Leave	11779
5	[5] NW-Military-Community Service	941
6	[6] NW-Unemployed	34668
8	[8] NW-But Sometimes Sec. Job	8664
9	[9] NW-work but past 7 days	3923
10	[10] NW-But Reg. Sec. Job	7756
11	[11] Working	380578
12	[12] Working But NW Past 7 Days	4489
-1	[-1] No Answer	33
-2	[-2] Does not apply	9
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

*Waves:* all

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

PGLFS provides a differentiation between “working” (Code 11–12) and “non-working” (Code 1–10), categories which are constant over all waves. Non-employment is subdivided further in order to make it possible to efficiently apply different labor market concepts in studying the data. To calculate this variable, the variables on employment status, age, maternity leave, second jobs, registration at the employment office, participation in paid work during the past 7 days and training status are used. Code (12) was added in 2000.

For respondents who have multiple status codes and different values for this variable, the following hierarchy was used to determine which of the values would play the determining role (increasing dominance):

value	label
11	working
1	non-working without further information
2	non-working, and older than 65
3	non-working, and currently in a training program
6	non-working, and registered unemployed
4	non-working, on maternity leave
5	non-working, in military/community service
9	non-working, but working past 7 days
10	non-working, but regular second job
8	non-working, but occasional second job
12	working, but non-working past 7 days

PGLFS supplements the variable PGEMPLST, which differentiates among persons who are employed. In 2016, the questionnaire for refugees contains a new labor market status category „vocational internship / betriebliches Praktikum“, which is integrated in the category 10 „education“ of the PGLFS variable.

For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

### pgjobch – Occupational Change

1	[1] Not Employed	272838
2	[2] Employed No Change	310716
3	[3] Employed No Info If Change	5966
4	[4] Employed With Change	61485
5	[5] First Time Employed	6897
-1	[-1] No Answer	30
-2	[-2] Does not apply	31
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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.

Hence, PGJOBCH indicates whether a respondent has changed jobs since the beginning of the previous year. The variable is calculated for all waves, and the codes are assigned independently of the respondent being a first-time or follow-up respondent.

The variable is also designed to identify respondents who have entered employment for the first time. Up to 1993, first-time respondents did not answer the question about job change. Therefore, for first-time respondents up to 1993, the variable was generated by using the information on the start date with the current employer and the respondent's age at entrance into his/her first job.

The variable is also designed to provide consistent longitudinal information on job changes. The 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 following:

1. For follow-up respondents generation as in case of a uniform questionnaire for all respondents to be pursued.
2. In the case of first-time respondents in M1 sample and for all respondents from M2 sample the rule of thumb is as in 2013.

In 2016 the respondents of the newly introduced refugee samples (M3 and M4, psample==17) were not asked whether they have changed jobs since the beginning of the previous year. The generation of PGJOBCH for these samples corresponds to the rule of thumb is as for the first-time respondent in the M1 sample with one modification. In step 4, the code (5) was assigned to respondents who are currently in a vocational training, were not occupied prior to their arrival in Germany, had neither a vocational training, nor a university degree and are in their first occupational relationship in Germany.

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

For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

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

1	[1] Worker/Employee (also mini job)	954
2	[2] Self-Employed, Free-Lance Professional	577
3	[3] Help In Family Business	120
4	[4] Other occupational position	1936
-1	[-1] No Answer	895



-2	[-2] Does not apply	22432
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	631049
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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	[1] Worker/Employee (also mini job)	69
2	[2] Self-Employed, Free-Lance Professional	102
3	[3] Help In Family Business	18
4	[4] Other occupational position	645
-1	[-1] No Answer	315
-2	[-2] Does not apply	25765
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	631049
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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	[1] Worker/Employee (also mini job)	9
2	[2] Self-Employed, Free-Lance Professional	23
3	[3] Help In Family Business	4
4	[4] Other occupational position	173
-1	[-1] No Answer	99
-2	[-2] Does not apply	26606
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	631049
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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)

? What is your current occupation? (from: [soep-core/soep-core-2017-pb-m345-erst/Q226:pber\[12711\]](#))

? What is your current position/occupation? (from: [soep-core/soep-core-2017-pe/52:pber\[4195\]](#))

0	[0] Soldiers	0
100	[100] Soldiers	1555
1000	[1000] Legislators, Senior Officials and Managers	0
1100	[1100] Legislators, Senior Officials and Managers	0
1110	[1110] Legislators, Senior Officials and Managers	136
1140	[1140] Senior Officials of Special-Interest Organisations	50
1141	[1141] Senior Officials of Political Party Organisations	9
1142	[1142] Senior Officials of Employers, Workers and Other Economic-Interest Organisations	256
1143	[1143] Senior Officials of Humanitarian and Other Special-Interest Organisations	4
1200	[1200] Corporate Managers	1304
1210	[1210] Directors and Chief Executives	2885
1220	[1220] Production and Operations Managers	0
1221	[1221] Production and Operations Managers in Agriculture, Hunting, Forestry and Fishing	18
1222	[1222] Production and Operations Managers in Manufacturing	1659
1223	[1223] Production and Operations Managers in Construction	0
...	(469 rows omitted)	346148
9213	[9213] Fishery, Hunting and Trapping Labourers	0
9300	[9300] Labourers in Mining, Construction, Manufacturing and Transport	0
9310	[9310] Mining and Construction Labourers	0
9311	[9311] Mining and Quarrying Labourers	49
9312	[9312] Construction and Maintenance Labourers: Roads, Dams and Similar Constructions	31
9313	[9313] Building Construction Laborer	1172
9320	[9320] Manufacturing Laborer	6655
9330	[9330] Transport Lab., Freight Handler	3182
-1	[-1] No Answer	2139

-2	[-2] Does not apply	290711
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

Waves: 1984-2017

Hartmann and Schütz (2002) provide detailed information on the conducted occupational coding. This result has been slightly modified to fit to the ISCO-88 version for European Union purposes (ISCO-88(COM)).

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))

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

? What is your current occupation? (from: [soep-core/soep-core-2017-pb-m345-erst/Q226:pber\[12711\]](#))

? What is your current position/occupation? (from: [soep-core/soep-core-2017-pe/52:pber\[4195\]](#))

110	[110] Commissioned Armed Forces Officers	49
210	[210] Non-Commissioned Armed Forces Officers	19
310	[310] Armed Forces Occupations, Other Ranks	179
1111	[1111] Legislators	26
1112	[1112] Senior Government Official	159
1113	[1113] Traditional Chiefs and Heads of Village	0
1114	[1114] Senior Officials of Special-Interest Organisations	51
1120	[1120] Managing Directors and Chief Executives	545
1211	[1211] Finance Managers	106
1212	[1212] Human Resource Managers	163
1213	[1213] Policy and Planning Managers	62
1219	[1219] Business Services and Administration Managers Not Elsewhere Classified	162
1221	[1221] Sales and Marketing Managers	453
1222	[1222] Advertising and Public Relations Managers	40
1223	[1223] Research and Development Managers	39
...	(413 rows omitted)	79232
9611	[9611] Garbage and Recycling Collectors	67
9612	[9612] Refuse Sorters	35
9613	[9613] Sweeper, Related Laborer	13
9621	[9621] Messengers, Package Deliverers and Luggage Porters	323
9622	[9622] Odd Job Persons	2
9623	[9623] Meter Readers and Vending-Machine Collectors	4
9624	[9624] Water and Firewood Collectors	0
9629	[9629] Elementary Workers Not Elsewhere Classified	84
-1	[-1] No Answer	750
-2	[-2] Does not apply	401678
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0

-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	173722

Waves: 2013–

Current occupation coded as ISCO-08. 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)

? What is your current occupation? (from: [soep-core/soep-core-2017-pb-m345-erst/Q226:pber\[12711\]](#))

? What is your current position/occupation? (from: [soep-core/soep-core-2017-pe/52:pber\[4195\]](#))

110	[110] Farmers, general	1812
111	[111] Fruit and vegetable farmers (non-horticultural)	5
112	[112] Arable farmers (special, permanent crops)	9
113	[113] Livestock farmers and pasture farmers	3
114	[114] Seed, crop producers, propagators (non-horticultural)	8
115	[115] Crop protectors	10
116	[116] Farmers and landlords	32
118	[118] Farmers and wine growers	12
120	[120] Wine growers, general	26
121	[121] Vine propagators	0
129	[129] Other wine growers	0
130	[130] Agricultural workers, general	106
131	[131] Agricultural supervisors	7
132	[132] Agricultural machinery drivers	58
133	[133] Vineyard workers	64
...	(2262 rows omitted)	361959
9832	[9832] Other employees (job-searching) with (as yet) undefined occupation	6
9911	[9911] Specialized professionals without further specification	95
9921	[9921] Homeworkers without further specification	151
9931	[9931] Forepersons, group leaders without further specification	353
9941	[9941] Persons doing community service without further specification	9
9951	[9951] Self-employed persons without further specification	364
9961	[9961] Consultancy, planning professionals without further specification	212
9971	[9971] Other employees without further specification	1598
-1	[-1] No Answer	718
-2	[-2] Does not apply	290346
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

Waves: 1984-2017

Current occupation coded as KldB92. 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)

? What is your current occupation? (from: [soep-core/soep-core-2017-pb-m345-erst/Q226:pber\[12711\]](#))

? What is your current position/occupation? (from: [soep-core/soep-core-2017-pe/52:pber\[4195\]](#))

1104	[1104] Officer	42
1203	[1203] Senior Non-Commissioned Officers and Higher	26
1302	[1302] Junior Non-Commissioned Officers	7
1402	[1402] Armed Forces Personnel in Other Ranks	177
11101	[11101] Occupations in Farming (without Specialisation)-Unskilled/Semiskilled Tasks	76
11102	[11102] Occupations in Farming (without Specialisation)-Skilled Tasks	264
11103	[11103] Occupations in Farming (without Specialisation)-Complex Tasks	11
11104	[11104] Occupations in Farming (without Specialisation)-Highly Complex Tasks	43
11113	[11113] Technical Occup. in Farming-Complex Tasks	3
11114	[11114] Technical Occup. in Farming-Highly Complex Tasks	0
11123	[11123] Agricultural Experts-Complex Tasks	0
11124	[11124] Agricultural Experts-High Complex Tasks	2
11132	[11132] Technical Laboratory Occup. in Agriculture-Skilled Tasks	0
11133	[11133] Technical Laboratory Occup. in Agriculture-Complex Tasks	0
11182	[11182] Occupations in Farming (with Specialisation, Not Elsewhere Classified)-Skilled Tasks	4
...	(1263 rows omitted)	81037
94623	[94623] Prop Designers-Complex Tasks	0
94693	[94693] Supervisors in Stage, Costume and Prop Design	1
94704	[94704] Occupations in Museums (without Specialisation)-Highly Complex Tasks	22
94712	[94712] Technical Occup. in Museums and Exhibitions-Skilled Tasks	3
94713	[94713] Technical Occup. in Museums and Exhibitions-Complex Tasks	0
94714	[94714] Technical Occup. in Museums and Exhibitions-Highly Complex Tasks	0
94724	[94724] Art Experts-Highly Complex Tasks	0
94794	[94794] Managers in Museum	0
-1	[-1] No Answer	839
-2	[-2] Does not apply	401684
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	173722

Waves: 2013–

Current occupation coded as KldB2010. 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](mailto:kwenzig@diw.de))

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

? What is your current occupation? (from: [soep-core/soep-core-2017-pb-m345-erst/Q226:pber\[12711\]](#))

? What is your current position/occupation? (from: soep-core/soep-core-2017-pe/52:pber[4195])

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
... (33 rows omitted)	287858
69	29527
70	3490
71	6135
74	1331
77	2933
78	941
79	368
82	1030
83	202
85	2310
87	838
88	3343
90	296
-1	1453
-2	180637

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 Leuw 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)

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

? What is your current occupation? (from: soep-core/soep-core-2017-pb-m345-erst/Q226:pber[12711])

? What is your current position/occupation? (from: soep-core/soep-core-2017-pe/52:pber[4195])

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
... (157 rows omitted)	415550
123.900001525879	1085
125.199996948242	255
132.100006103516	3298
135.699996948242	1775
138.199996948242	390
138.899993896484	202
139.800003051758	394
145.699996948242	2220
152.5	2365
153.5	162
191.300003051758	3343
207.199996948242	614
216	781
-1	1536
-2	180752

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 the Federal Republic of Germany. MPS is assigned based on the German Federal Statistical Office's occupational classification of 1992 (PGKLDB92). 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 Value (Std. Internat. Occupational Prestige Scale), Treim

? What is your current occupation? (from: [soep-core/soep-core-2017-pb-m345-erst/Q226:pber\[12711\]](#))

? What is your current position/occupation? (from: [soep-core/soep-core-2017-pe/52:pber\[4195\]](#))

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
... (34 rows omitted)	387177
64	830
65	2558
66	3651
67	1088
68	31
69	697
70	6807
71	1030
72	1829
73	1529
75	194
76	296
78	5811
-1	1453
-2	180637

Waves: 1984-2017

This variable gives the occupational prestige score index for all employed persons. PGSIOPS88 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 Fritzsche and Wirth (2001).

Please also see occupational prestige scores (PGMPS92), occupational status (PGISEI88), and occupational class (PGEGP88).

*Fritzsche, 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)

### **pgexp88** – Last Reached Exp Value (Erikson, Goldthorpe, Portocarero)

? What is your current occupation? (from: [soep-core/soep-core-2017-pb-m345-erst/Q226:pber\[12711\]](#))

? What is your current position/occupation? (from: [soep-core/soep-core-2017-pe/52:pber\[4195\]](#))

1	[1] [I] Higher Managerial and Professional Workers	56915
2	[2] [II] Lower Managerial and Professional Workers	102249
3	[3] [IIIa] Routine Clerical Work	59359
4	[4] [IIIb] Routine Service and Sales Work	61426
5	[5] [IVa] Small Self-Employed With Employees	6839
6	[6] [IVb] Small Self-Employed Without Employees	13458



7	[7] [V] Manual Supervisors	0
8	[8] [VI] Skilled Manual Workers	88770
9	[9] [VIIa] Semi- and Unskilled Manual Workers	77623
10	[10] [VIIb] Agricultural Labour	6553
11	[11] [IVc] Self-Employed Farmers	2681
-1	[-1] No Answer	1453
-2	[-2] Does not apply	180637
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

#### Waves: 1984-2017

This variable gives the occupational class for all employed persons. PGEGP88 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 PGEGP88 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, PGMPS92).

John Hendrickx, 2002. "ISCO: 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	[1] Yes	199180
2	[2] No	126084
3	[3] In Training	21475
4	[4] Has No Job Training	25238
-1	[-1] No Answer	13538
-2	[-2] Does not apply	272448
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0

-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

*Waves:* all

This variable is designed to offer annual data on all employed persons, indicating whether they are working in the occupation they were trained for.

*For more information, contact:* Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

### pgausb – Required Training For Job

---

1	[1] No Training	17360
2	[2] Intro. To Job	48361
3	[3] On-The-Job Training	28275
4	[4] Courses	14725
5	[5] Vocational Training	180887
6	[6] Technical School, Engineering (East) 90-96	2002
7	[7] Technical College or University, until 98	10446
8	[8] Technical College since 99	24494
9	[9] University since 99	35750
-1	[-1] No Answer	15710
-2	[-2] Does not apply	279953
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

*Waves:* all

This variable is designed to provide annual data on required job training for all employed persons. The variable is generated using questions on required formal education and required on-the-job-training which are categorized into seven independent variables with 0/1 coding. Out of these, the highest available level of required training is used for the generation of the status variable.

The answer option “completed technical school” was only used from 1990 to 1993 in the East German version of the questionnaire. Since not all employed people are asked the question about required training every year, the value (6) of the variable PGAUSB is valid up to 1996. The code (-2) is assigned to all non-employed persons and also includes persons in occupational training, in occupational retraining programs, and those doing an internship at the time of the survey.

*For more information, contact:* Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

### pgerwzeit – Length Of Time With Firm

---

*Waves:* all

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

The variable provides consistent longitudinal information on the length of time with the same employer. Data that show longitudinal inconsistencies are corrected, if it can be done under reasonable assumptions.

1. In case of no job change, the information on the start date with the current employer given in the earliest interview available is treated as dominant and carried forward to the subsequent years.
2. In case of a job change between firms, the information on the start of the current position is used and carried forward to the subsequent years.
3. Up to wave Z (2009), a respondent who starts working again after a period of non-employment is assumed to have returned to the former employer if the indicated start date with the current employer was before the previous interview date. In this case, the start date with the current employer given in the previous interview is treated as dominant. Otherwise, the present information on the start date with the current employer is used and carried forward to the subsequent years. For respondents who are assumed to have returned to their former employer, the full length of time with the firm is calculated. There is no deduction for the time during which the respondent was not employed.
4. Since wave BA (2010), there is a modified answer category in the questionnaire which indicates that a respondent returns to his/her former employer after a period of non-employment. If a respondent indicates to have started working again at a former employer, the present information on the start date with the current employer is used and carried forward to the subsequent years. Unlike before wave BA (2010), the present information is treated as dominant even if the indicated start date with the current employer was before the previous interview date. Hence, the full length of time with the firm is calculated, and there is no deduction for the time during which the respondent was not employed or employed in another firm.
5. The length of time with the firm is also provided for the East German sample since its start in 1990. Due to the massive restructuring of the economy that took place in East Germany after reunification, this variable should be dealt with cautiously in the first transition years.

Both monthly and annual information is used in the variables and rounded off as length of time in years (with months in decimal form).

*For more information, contact:* Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

### pgtatzeit – Actual Work Time Per Week

*Waves:* all

This variable provides data on actual weekly working hours (including overtime) for all persons employed at the time of the survey (including the self-employed). The data are obtained by asking respondents how many hours they work on average per week.

For implausible answers (actual weekly working hours of more than 80 per week), we assign the value (-3). The variable is rounded off and gives the number of working hours as a decimal number.

Please also see PGVEBZEIT and PGUEBSTD.

*For more information, contact:* Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

**pgvebzeit** – Agreed Upon Work Time Per Week

---

*Waves:* all

This variable provides data on agreed weekly working hours. The variable takes into account only those persons who were in dependent employment (not self-employed) at the time of the survey.

The value (-2) is assigned to non-employed people, employees without set hours and to self-employed people, including self-employed farmers, freelancers, and other self-employed persons. If persons helping out in family businesses report agreed weekly working hours, we assign a non-missing value.

For implausible answers (agreed weekly working time of more than 80 hours per week) we assign the value (-3). The variable is rounded off and gives the number of working hours as a decimal number.

Please also see PGTATZEIT and PGUEBSTD.

*For more information, contact:* Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

**pguebstd** – Overtime Per Week

---

0	[0] No Overtime Last Month	180645
-1	[-1] No Answer	13400
-2	[-2] Does not apply	311504
-3	[-3] Answer improbable	132
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

*Waves:* all

This variable provides data on overtime per week for all persons in dependent employment at the time of the survey. The data are obtained by asking respondents how many overtime hours they worked in the month before the survey. The number of monthly overtime hours is then converted into weekly overtime by dividing the number given by 4.3. Since 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. The variable PGUEBSTD for these years was therefore generated using the difference between average actual weekly working hours and agreed weekly working hours.

The value (-2) is assigned to non-employed people, employees without set hours and to self-employed people, including self-employed farmers, freelancers, and other self-employed persons. If persons helping out in family businesses report overtime hours, we assign a non-missing value.

For implausible answers (agreed-upon weekly working time or actual weekly working time of more than 80 hours per week AND weekly overtime of more than 10 hours) we assign the value (-3).

The variable is rounded off and gives the number of overtime hours as a decimal number.

Please also see PGVEBZEIT and PGTATZEIT.

*For more information, contact:* Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

**pgoeffd** – Civil Service

---

1	[1] Yes	89837
2	[2] No	274635
-1	[-1] No Answer	21043
-2	[-2] Does not apply	272448
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

Waves: all

Status variable: This variable is designed to provide annual data on employment in the civil service for all employed persons.

For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, afedorets@diw.de)

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

**?** In which economic sector/branch of industry/service area is the company or institution that you work for? (*from: soep-core/soep-core-2017-pb-m345-erst/Q229:pbra[12713]*)

**?** What sector of business or industry is your company or institution active in for the most part? (*from: soep-core/soep-core-2017-pe/55:pbra[4197]*)

1	[1] Agriculture, Hunting, Related Service Activities	5972
2	[2] Forestry, Logging, Related Service activities	550
5	[5] Fishing, Operation Of Fish Hatcheries And Fish Farms	33
10	[10] Mining Of Coal And Lignite; Extraction Of Peat	1055
11	[11] Extraction Of Crude Petroleum And Natural Gas	109
12	[12] Mining Of Uranium And Thorium Ores	0
13	[13] Mining Of Metal Ores	14
14	[14] Other Mining And Quarrying	215
15	[15] Manuf Food Products And Beverages	7448
16	[16] Manuf Tobacco Products	87
17	[17] Manuf Textiles	2605
18	[18] Manuf Wearing Apparel; Dressing And Dyeing Of Fur	933
19	[19] Tanning,Dressing Of Leather; Manuf luggage, Footwear	434
20	[20] Manuf Wood Products, Except Furniture	1497
21	[21] Manuf Pulp, Paper And Paper Products	1429
...	(39 rows omitted)	306989
90	[90] Sewage And Refuse Disposal, Sanitation And Related	1376
91	[91] Activities Of Membership Organizations NEC.	4232
92	[92] Recreational, Cultural And Sporting Activities	5345
93	[93] Other Service Activities	3443
95	[95] Private Households With Employed Persons	1513
96	[96] Industry - NEC	0
97	[97] Handcraft, Trade - NEC	0
99	[99] Extra-territorial Organizations And Bodies	260
-1	[-1] No Answer	2845
-2	[-2] Does not apply	309579
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0

-6	[6] Version of questionnaire with modified filtering	0
-8	[8] Question this year not part of Survey program	0

*Waves: 1984-2017*

This variable is designed to provide annual data on the industry of economic activity for all employed persons according to the Statistical Classification of Economic Activities in the European Community (Nomenclature des statistiques des activités économiques de la Communauté européenne - NACE Rev. 1.1 which is on this level similar to Rev. 1). Respondents answer the question in their own words regarding the industry in which they are currently working. In order to facilitate international comparability, the European industry standard classification system is used by Kantar Public using Level 2 (Divisions) of the classification system. The classification has been documented in Hartmann/Schütz 2002.

The codes in NACE Rev.1 also correspond to ISIC Rev.3 (International Standard Classification of All Economic Activities).

*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	[1] Crop and animal production, hunting and related service activities	843
2	[2] Forestry and logging	126
3	[3] Fishing and aquaculture	10
5	[5] Mining of coal and lignite	50
6	[6] Extraction Of Crude Petroleum And Natural Gas	30
7	[7] Mining Of Metal Ores	0
8	[8] Other Mining And Quarrying	27
9	[9] Mining support service activities	1
10	[10] Manufacture of food products	1664
11	[11] Manufacture of beverages	133
12	[12] Manuf Tobacco Products	8
13	[13] Manuf Textiles	266
14	[14] Manuf Wearing Apparel; Dressing And Dyeing Of Fur	140
15	[15] Manufacture of leather and related products	39
16	[16] Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	308
...	(65 rows omitted)	74434
92	[92] Gambling and betting activities	74
93	[93] Sports activities and amusement and recreation activities	376
94	[94] Activities of membership organisations	996
95	[95] Repair of computers and personal and household goods	73
96	[96] Other personal service activities	806
97	[97] Private Households With Employed Persons	448
98	[98] Undifferentiated goods- and services-producing activities of private households for own use	0
99	[99] Extra-territorial Organizations And Bodies	38
-1	[-1] No Answer	1004
-2	[-2] Does not apply	416174
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0

-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	159895

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

### **pgbetr** – Size of the Company

1	[1] Lt 5	33212
2	[2] Ge 5 Lt 10	15898
3	[3] Ge 11 Lt 20	13637
4	[4] Until 90: Lt 20	9940
5	[5] 91-04: Ge 5 Lt 20	25518
6	[6] Ge 20 Lt 100	46055
7	[7] Ge 100 Lt 200	21714
8	[8] Until 98: Ge 20 Lt 200	30259
9	[9] Ge 200 Lt 2000	74807
10	[10] Ge 2000	82859
11	[11] Self-Employed Without Coworkers	13144
-1	[-1] No Answer	19058
-2	[-2] Does not apply	271862
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

*Waves: all*

This variable provides data on the total size of the company (number of employees at all company locations).

Please pay attention to special codes 4, 5, and 8! These codes were necessary due to the differentiation of items for small and medium-sized companies over the years. In the years 1991, 1999, and 2005, respondents were asked about company size in a more detailed form, so when the data were recalculated for the years 1984–90, the original codes (1)–(5) were changed to (4)(8)(9)(10) and (11). For the years 1999 to 2004, Code (8) was differentiated into (6) and (7). From 2005 on, the prior category “5 to 20 employees” (5) has been split into the two categories “5 to 10 employees” (2) and “11 to 20 employees” (3).

In 2012, the questionnaire provides a one-time-only information on the size of the local establishment in addition to the size of the entire company (PGBETR). The enriched questionnaire revealed that in previous interviews, persons have mistakenly provided information on the local establishment size instead of the entire company size, especially if their entire



company had 2000 and more employees. Due to the importance of longitudinal consistency, these persons 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 persons without a job change were modified if:

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

Those person's values of the entire company in 2012 were replaced by their values of the local establishment size in 2012.

Please also see PGALLBET for a broader categorization of the firm size, which is appropriate for analyses that include all sample years.

*For more information, contact:* Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

### pgallbet – Core Category Size Of The Company

---

1	[1] LT 20	98205
2	[2] GE 20 LT 200	98028
3	[3] GE 200 LT 2000	74807
4	[4] GE 2000	82859
5	[5] Self-Employed Without Coworkers	13144
-1	[-1] No Answer	19058
-2	[-2] Does not apply	271862
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

*Waves:* all

This variable is designed to provide annual data on the core size category of the company for all employed persons. Since respondents were asked about company size in more detailed form in the years 1991, 1999, and 2005 (see also 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 values of variables BETR84 to BETR90 and offers a variable that is consistent across all waves.

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

*For more information, contact:* Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

## 8 Last Occupation

### pgjobend – Reasons for occupational change

---

1	[1] Terminated by employer	10091
2	[2] Limited Term Employment Relationship	1834
3	[3] Education, training completed	769
4	[4] Own resignation	14738
5	[5] Mutual termination	3572



6	[6] Employee requested transfer	448
7	[7] Company transferred employee	492
8	[8] Ended self-employment	1560
9	[9] Temporary contract or education/training completed	6060
10	[10] Took early retirement	552
11	[11] Company closed down	3309
12	[12] Old-age pension	3032
13	[13] Leave of absence/sabbatical (1999-2010)	1786
14	[14] Leave, maternity leave and parental leave (1991-1998), since 2011	2537
15	[15] Other incl. early retirement, company closed, old-age pension, leave of absence/sabbatical (1985-1986)	173
16	[16] Other incl. company closed, old-age pension, leave of absence/sabbatical (1987-1990)	128
17	[17] Other incl. mutual termination (1991-1998)	687
-1	[-1] No Answer	2828
-2	[-2] Does not apply	593390
-3	[-3] Answer improbable	19
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	9958
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

#### Waves: 1985–

This variable is designed to offer annual data on reasons for an occupational change for all formerly employed persons, persons with a job change or persons on leave. For years 1985–1990 also persons who changed positions in the same company are considered. Only persons with valid dates for an occupational or positional change are included. Likewise to the questionnaire the variable offers data from interview date to interview date not from one year to the following. Respondents are asked about their annual and possibly same occupational change in two consecutive interviews, duplicate answers are therefore considered only once and the older statement is dominant. If a respondent stated a job termination in the current interview which was before the interview date in the previous year but didn't reported this in the previous interview this termination has been counted for the current interview. For years 1985 up to 1998 every given reason was coded as separate variable with variable values "Yes" (1) and "Does not apply" (-2), which resulted in up to 13 different variables. Since 1999 all given reasons have been collected in one single variable with diverse values. If respondent states more than one reason for job termination, the first of the stated reasons is dominant. The year 1990 is specific due to introduction of sample for East Germany. Since the questionnaire did not contain the information on reasons for end of the job all the observations in Sample C in year 1990 obtained value (-5).

Please pay attention to special codes (15), (16), and (17)! These codes were necessary due to the variety of the given values over the years. In any years respondents were asked about reasons for change with more or less given answers and from years 1985–1998 also the answer "Other" was possible. While all explicit reasons have been recoded to uniform values, the answer "Other" then in some years includes reasons for which in other years was separately asked for: "Other" was coded (15) for years 1985 and 1986, (16) for years 1987–1990 and (17) for years 1991–1998.

For years 1991–1998 and 2011–2012 there is a variable value "Leave, maternity leave and parental leave" whereas for 1999–2010 the given reason covered only "Leave of absence/sabbatical". Note that codes (2) and (3) for years 1985–1998 have been merged to code (9) since 1999. In 2013 the respondents of the newly introduced migration sample were not asked about

an occupational change, but in 2014 they were. So information on jobend for the migration sample in 2013 were taken from the questionnaire in 2014.

Since 2006 youth questionnaires have been embodied in survey tools. Nevertheless, the questionnaires do not contain the information on jobend. Therefore, in order to account for the persons represented by the youth questionnaires the negative value (-5) has been introduced starting from year 2006.

*For more information, contact:* Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

## 9 Employment History

### pgexpft – Working Experience Full-Time Employment

*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 gives the length of time in years with months in decimal form.

If there is no monthly calendar data available in a given year of a respondent's career, the annual data from PBIOSPE is used for that year. In the most current wave the variable only uses up-to-date information from the newly answered Biography Questionnaires. If the year in which a spell started and ended is the same, and if there is no monthly data, a spell of 0.5 years is assumed. Persons without annual data (not contained in PBIOSPE) are only assigned a non-missing value for this variable if they joined SOEP by the age of 18 and if there is calendar data on them in ARTKALEN.

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

Please also see PGEXPPT and PGEXPUE.

*For more information, contact:* Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

### pgexppt – Working Experience Part-Time Employment

0	379483
0.100000001490116	4192
0.200000002980232	4858
0.300000011920929	8101
0.400000005960464	3720
0.5	13582
0.600000023841858	3736
0.699999988079071	3251
0.800000011920929	4916
0.899999976158142	2565
1	20511
1.10000002384186	2569
1.20000004768372	2644
1.29999995231628	5236
1.39999997615814	2413

...	(449 rows omitted)	180149
47.2000007629395		1
47.2999992370605		3
47.7000007629395		1
48		4
48.0999984741211		1
48.2999992370605		1
48.4000015258789		1
48.7000007629395		1
49		1
49.2000007629395		1
49.7000007629395		1
50		18
50.7000007629395		1
51		1
-1		16001

#### 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 gives the length of time in years with months in decimal form.

If there is no monthly calendar data available in a given year of a respondent's career, the annual data from PBIOSPE is used for that year. In the most current wave the variable only uses up-to-date information from the newly answered Biography Questionnaires. If the year in which a spell started and ended is the same, and if there is no monthly data, a spell of 0.5 years is assumed. Persons without annual data (not contained in PBIOSPE) are only assigned a non-missing value for this variable if they joined SOEP by the age of 18 and if there is calendar data on them in ARTKALEN.

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

Please also see PGEXPFT and PGEXPUE.

For more information, contact: Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

#### pgexpue – Unemployment Experience

0	432425
0.100000001490116	9410
0.200000002980232	8127
0.300000011920929	11749
0.400000005960464	5497
0.5	29487
0.600000023841858	5295
0.699999988079071	5066
0.800000011920929	6915
0.899999976158142	3623

1	19650
1.10000002384186	4190
1.20000004768372	3600
1.29999995231628	4448
1.39999997615814	2767
... (250 rows omitted)	89647
29	3
29.2999992370605	8
29.5	1
30	1
31	3
32	4
33	5
34	2
34.2999992370605	1
35	2
36	1
37	4
38	30
40	1
-1	16001

*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 gives the length of time in years with months in decimal form.

If there is no monthly calendar data available on a given year in a respondent's career, the annual data from PBIOSPE is used for that year. In the most current wave the variable only uses up-to-date information from the newly answered Biography Questionnaires. If the year in which a spell started and ended is the same, and if there is no monthly data, a spell of 0.5 years is assumed. Persons without annual data (not contained in PBIOSPE) are only assigned a non-missing value for this variable if they joined SOEP by the age of 18 and if there is calendar data on them in ARTKALEN.

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

Please also see PGEXPFT and PGEXPPT.

*For more information, contact:* Alexandra Fedorets (Tel. +49-30-89789-321, [afedorets@diw.de](mailto:afedorets@diw.de))

## 10 Volunteering

### pghonor1 – Third Voluntary Work

1	[1] Yes	2911
2	[2] No	1524
-1	[-1] No Answer	47

-2	[-2] Does not apply	22432
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	631049
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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)

### pgonor2 – Second Voluntary Work

---

1	[1] Yes	961
2	[2] No	171
-1	[-1] No Answer	17
-2	[-2] Does not apply	25765
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	631049
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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)

### pgonor3 – First Voluntary Work

---

1	[1] Yes	267
2	[2] No	34
-1	[-1] No Answer	7
-2	[-2] Does not apply	26606
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	631049
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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	[0] in school	16268
1	[1] inadequately	23461
2	[2] general elementary	101897
3	[3] middle vocational	309103
4	[4] vocational + Abi	35191
5	[5] higher vocational	42574
6	[6] higher education	115963
-1	[-1] No Answer	13506
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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	[0] in school	6511
1	[1] Primary education	7893
2	[2] Lower secondary education	28291
3	[3] Upper secondary education	104808
4	[4] Post-secondary non-tertiary education	18054
5	[5] Short-cycle tertiary education	9784
6	[6] Bachelors or equivalent level	33806
7	[7] Masters or equivalent level	17334
8	[8] Doctoral or equivalent level	1923
-1	[-1] No Answer	6947
-2	[-2] Does not apply	422612
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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.

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

### pgcasmin – CASMIN Classification

---

0	[0] (0) In School	16171
1	[1] (1a) Inadequately Completed	26233
2	[2] (1b) General Elementary School	80648
3	[3] (1c) Basic Vocational Qualification	168470
4	[4] (2b) Intermediate General Qualification	22919
5	[5] (2a) Intermediate Vocational	136176
6	[6] (2c_gen) General Maturity Certificate	28295
7	[7] (2c_voc) Vocational Maturity Certificate	45929
8	[8] (3a) Lower Tertiary Education	42789
9	[9] (3b) Higher Tertiary Education	73174
-1	[-1] No Answer	17159
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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	23557
8.5	2084
9	82709
10	32493
10.5	125376
11	39029
11.5	97152
12	55171
13	33929
13.5	7380
14	8633
14.5	14148

15	30131
16	15192
17	1956
18	55696
-1	18132
-2	15195

*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	[1] Secondary School Degree	217501
2	[2] Intermediate School Degree	175859
3	[3] Technical School Degree	29321
4	[4] Upper Secondary Degree	112396
5	[5] Other Degree	61771
6	[6] Dropout, No School Degree	25362
7	[7] No School Degree Yet	16986
8	[8] No School Attended	813
-1	[-1] No Answer	17954
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0



-8 [-8] Question this year not part of Survey program 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	[1] Apprenticeship	258173
2	[2] Vocational School	49450
3	[3] Health Care School	8598
4	[4] Technical School	37108
5	[5] Civil Service Training	15171
6	[6] Other degree	24580
7	[7] Completed Vocational Training/Education in Germany	1631
-1	[-1] No Answer	8213
-2	[-2] Does not apply	255039
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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	[1] Technical College	32787
2	[2] University, Technical College	55162
3	[3] College Not In Germany	8112
4	[4] Engineering, Technical School (East)	10852
5	[5] University (East)	7004
6	[6] graduation, state doctorate	1628
7	[7] graduation, state doctorate (foreign country, east)	299
8	[8] institution of higher education (youth)	0

9	[9] Dual Studies, University of Cooperative Education	110
10	[10] Other Colleges (since 2014)	60
-1	[-1] No Answer	8175
-2	[-2] Does not apply	533774
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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)

### pgpbil03 – No Vocational Degree

---

1	[1] No Vocation Degree	146103
2	[2] Apprenticeship	25060
3	[3] University	19812
-1	[-1] No Answer	8236
-2	[-2] Does not apply	458752
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

*Waves:* all

In connection with the question about vocational degrees (PGPBIL01 and PGPBIL02), 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	[1] 8th Grade Completed	35579
2	[2] 10th Grade Completed	60938
3	[3] College Entrance Exam	20671
4	[4] Other Degree	1215
5	[5] Dropout, No School Degree	1406
6	[6] No School Degree Yet	0

-1	[-1] No Answer	15
-2	[-2] Does not apply	473909
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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	[1] Vocational Training	61574
2	[2] Master Craftsman	6530
3	[3] Engineering, Technical Degree	15272
4	[4] Other degree	1079
-1	[-1] No Answer	0
-2	[-2] Does not apply	509278
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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	[1] School, No Degree	17854
2	[2] School, With Degree	35633
3	[3] Vocational Extension School	25865
4	[4] School Leaving Degree[Sbil] Acquired Abroad	12
5	[5] Certificate from a different school	277
-1	[-1] No Answer	197

-2	[-2] Does not apply	578125
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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)

### pgpbbila – Vocational Degree Outside Germany

---

1	[1] On-The-Job Training	2567
2	[2] Vocational Training	3623
3	[3] Vocational School	3386
4	[4] College	1846
5	[5] Other	653
6	[6] Vocational Degree[Bbil01] Acquired Abroad	24
7	[7] College Education[Bbil02] Acquired Abroad	53
8	[8] Completed Vocational Training/Education Other Country	866
9	[9] graduation, state doctorate (foreign country)	23
11	[11] With Certificate, On-The-Job Training	1486
12	[12] With Certificate, Vocational Training	5959
13	[13] With Certificate, Vocational School	7560
14	[14] With Certificate, University	10598
15	[15] With Certificate, Other	1246
16	[16] With Certificate, Vocational Degree[bbil01] Acquired Abroad	46
17	[17] With Certificate, College Education[Bbil02] Acquired Abroad	83
18	[18] With Certificate, Completed Vocational Training/Education Other Country	0
19	[19] With Certificate, PhD (Other Country)	282
-1	[-1] No Answer	144
-2	[-2] Does not apply	614847
-3	[-3] Answer improbable	2671
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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	[1] Sprach- und Kulturwissenschaften allgemein	396
2	[2] Protestant Theology	608
3	[3] Catholic Theology	27

4	[4] Philosophy	259
5	[5] History	366
6	[6] Bibliothekswissenschaft, Dokumentation, Publizistik	195
7	[7] Allgemeine und vergleichende Literatur und Sprachwissenschaft	636
8	[8] Altphilologie (klass. Philologie), Neugriechisch	133
9	[9] Germanistik (Deutsch, germanische Sprachen ohne Anglistik)	1305
10	[10] English Studies	454
11	[11] Romance Philology	195
12	[12] Slavonic Studies	75
13	[13] Aussereuropäische Sprach- und Kulturwissenschaften	125
14	[14] Cultural Studies	90
15	[15] Psychology	776
...	(36 rows omitted)	45224
68	[68] Civil Engineering	1739
69	[69] Surveying and Mapping	196
74	[74] Art, Aesthetics	359
75	[75] Fine Arts	74
76	[76] Design	581
77	[77] Darstellende Kunst, Film und Fernsehen, Theaterwissenschaft	168
78	[78] Music, Musicology	613
83	[83] Outside the structure of the university system	230
-1	[-1] No Answer	87217
-2	[-2] Does not apply	514820
-3	[-3] Answer improbable	1102
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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
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

11	[11] Magister	1606
12	[12] Diplom (University)	14406
13	[13] Bachelor	988
14	[14] Master	804
15	[15] 1st State Examination	1906
16	[16] Other state examination	1986
21	[21] Diplom (at technical college, technical college for administration)	10318
22	[22] Bachelor (at technical college, technical college for administration)	697
23	[23] Master (at technical college, technical college for administration)	149
31	[31] Teacher training, BA, MA at elementary, lower secondary schools/primary level	1746
32	[32] Teacher training, BA, MA at 2ndary level 1/elementary schools/primary level	102
33	[33] Teacher training, BA, MA at intermediate scndry schools/scndry level I	791
34	[34] Teacher training, BA, MA at secondary level II and I	49
35	[35] Teacher training, BA, MA at academic 2ndry schools, 2ndry lev 2, genrl school	1127
36	[36] Teacher training, BA, MA at special needs schools	464

37	[37] Teacher training, BA, MA at vocational schools	196
38	[38] Teacher training, other	4833
41	[41] Academic degree in the arts	210
42	[42] Doctorate	2353
43	[43] Post-doctoral dissertation (Habilitation)	62
44	[44] Other degree	2007
98	[98] Not categorizable	600
-1	[-1] No Answer	95366
-2	[-2] Does not apply	515001
-3	[-3] Answer improbable	196
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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	[110] Farmers, general	1423
111	[111] Fruit and vegetable farmers (non-horticultural)	8
112	[112] Arable farmers (special, permanent crops)	28
113	[113] Livestock farmers and pasture farmers	0
114	[114] Seed, crop producers, propagators (non-horticultural)	7
115	[115] Crop protectors	28
116	[116] Farmers and landlords	54
118	[118] Farmers and wine growers	11
120	[120] Wine growers, general	30
121	[121] Vine propagators	0
129	[129] Other wine growers	0
130	[130] Agricultural workers, general	367
131	[131] Agricultural supervisors	0
132	[132] Agricultural machinery drivers	0
133	[133] Vineyard workers	0
...	(2262 rows omitted)	139117
9832	[9832] Other employees (job-searching) with (as yet) undefined occupation	0
9911	[9911] Specialized professionals without further specification	119
9921	[9921] Homeworkers without further specification	0
9931	[9931] Forepersons, group leaders without further specification	0
9941	[9941] Persons doing community service without further specification	0
9951	[9951] Self-employed persons without further specification	0
9961	[9961] Consultancy, planning professionals without further specification	7
9971	[9971] Other employees without further specification	10104
-1	[-1] No Answer	4336



-2	[-2] Does not apply	502324
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

#### Waves: -1985

The variable is designed to provide information on the occupation of vocational training which adds details to the information recorded in the variable 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 ("Beamtenausbildung"). 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-oekonomischen 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	[110] Farmers, general	43
111	[111] Fruit and vegetable farmers (non-horticultural)	0
112	[112] Arable farmers (special, permanent crops)	0
113	[113] Livestock farmers and pasture farmers	0
114	[114] Seed, crop producers, propagators (non-horticultural)	0
115	[115] Crop protectors	0
116	[116] Farmers and landlords	8
118	[118] Farmers and wine growers	0
120	[120] Wine growers, general	0
121	[121] Vine propagators	0
129	[129] Other wine growers	0
130	[130] Agricultural workers, general	41
131	[131] Agricultural supervisors	0
132	[132] Agricultural machinery drivers	0
133	[133] Vineyard workers	0
...	(2262 rows omitted)	23595
9832	[9832] Other employees (job-searching) with (as yet) undefined occupation	0
9911	[9911] Specialized professionals without further specification	51
9921	[9921] Homeworkers without further specification	0
9931	[9931] Forepersons, group leaders without further specification	0
9941	[9941] Persons doing community service without further specification	0
9951	[9951] Self-employed persons without further specification	0
9961	[9961] Consultancy, planning professionals without further specification	0
9971	[9971] Other employees without further specification	4889
-1	[-1] No Answer	1323
-2	[-2] Does not apply	628013
-3	[-3] Answer improbable	0

-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

*Waves: -1985*

The variable is designed to provide information on the occupation of full-time school based vocational training (e.g., Berufsfachschule, Schule des Gesundheitswesens, Handelsschule). See the description of variable 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	[110] Farmers, general	350
111	[111] Fruit and vegetable farmers (non-horticultural)	0
112	[112] Arable farmers (special, permanent crops)	41
113	[113] Livestock farmers and pasture farmers	0
114	[114] Seed, crop producers, propagators (non-horticultural)	0
115	[115] Crop protectors	0
116	[116] Farmers and landlords	0
118	[118] Farmers and wine growers	0
120	[120] Wine growers, general	17
121	[121] Vine propagators	0
129	[129] Other wine growers	0
130	[130] Agricultural workers, general	14
131	[131] Agricultural supervisors	0
132	[132] Agricultural machinery drivers	0
133	[133] Vineyard workers	0
...	(2262 rows omitted)	15456
9832	[9832] Other employees (job-searching) with (as yet) undefined occupation	0
9911	[9911] Specialized professionals without further specification	0
9921	[9921] Homeworkers without further specification	0
9931	[9931] Forepersons, group leaders without further specification	0
9941	[9941] Persons doing community service without further specification	0
9951	[9951] Self-employed persons without further specification	0
9961	[9961] Consultancy, planning professionals without further specification	0
9971	[9971] Other employees without further specification	3130
-1	[-1] No Answer	922
-2	[-2] Does not apply	638033
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

*Waves: -1985*

The variable is designed to provide information on the occupation of higher level vocational training (e.g., Meister, Techniker). See the description of variable 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	[110] Farmers, general	2
111	[111] Fruit and vegetable farmers (non-horticultural)	0
112	[112] Arable farmers (special, permanent crops)	0
113	[113] Livestock farmers and pasture farmers	0
114	[114] Seed, crop producers, propagators (non-horticultural)	0
115	[115] Crop protectors	0
116	[116] Farmers and landlords	0
118	[118] Farmers and wine growers	0
120	[120] Wine growers, general	0
121	[121] Vine propagators	0
129	[129] Other wine growers	0
130	[130] Agricultural workers, general	0
131	[131] Agricultural supervisors	0
132	[132] Agricultural machinery drivers	0
133	[133] Vineyard workers	0
...	(2262 rows omitted)	5436
9832	[9832] Other employees (job-searching) with (as yet) undefined occupation	0
9911	[9911] Specialized professionals without further specification	0
9921	[9921] Homeworkers without further specification	0
9931	[9931] Forepersons, group leaders without further specification	0
9941	[9941] Persons doing community service without further specification	0
9951	[9951] Self-employed persons without further specification	0
9961	[9961] Consultancy, planning professionals without further specification	0
9971	[9971] Other employees without further specification	1853
-1	[-1] No Answer	397
-2	[-2] Does not apply	650275
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	0

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	[1] Individual Questionnaire	17451
2	[2] Gap Questionnaire (temporary drop-outs)	633
3	[3] Biographical Questionnaire	163678
4	[4] Various Sources	0
-1	[-1] No Answer	0

-2	[-2] Does not apply	476201
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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	[0] Inconsistent educational information since last year	521055
1	[1] Inconsistent educational information since last survey	3212
2	[2] Inconsistent educational information since last year	3910
-1	[-1] No Answer	0
-2	[-2] Does not apply	129786
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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)

### pgbilztev – Change in Education, total observed period

---

0	[0] Consistent educational information	493895
1	[1] Inconsistent educational decline	53516
2	[2] Inconsistent educational increase	65986
3	[3] Inconsistent educational decline and increase	1553
-1	[-1] No Answer	0
-2	[-2] Does not apply	43013
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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	[1] Yes	241
2	[2] No	4088
-1	[-1] No Answer	136
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	25405
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	628093

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

---

	628093
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	[0] Not Categorized	57
1	[1] ECO_Ecological Parties	0
2	[2] COM_Socialist Parties	7
3	[3] SOC_Social Democratic Parties	6
4	[4] LIB_Liberal Parties	3
5	[5] CHR_Christian Democratic Parties	0
6	[6] CON_Conservative Parties	10
7	[7] NAT_Nationalist Parties	23
8	[8] AGR_Agrarian Parties	0
9	[9] ETH_Ethnic and Regional Parties	22
10	[10] SIP_Special Issue Party	0
11	[11] ISP_Islamic Parties	7
-1	[-1] No Answer	106
-2	[-2] Does not apply	4224
-3	[-3] Answer improbable	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	25405
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	628093

## 13 Information on the Interview

### pgmonth – Month Of Interview

---

1	[1] January	42188
---	-------------	-------



2	[2] February	160254
3	[3] March	174196
4	[4] April	99898
5	[5] May	57378
6	[6] June	44605
7	[7] July	30741
8	[8] August	20036
9	[9] September	14177
10	[10] October	8545
11	[11] November	3842
12	[12] December	2054
-1	[-1] No Answer	0
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	49
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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

---

2017	31786
2018	1336
-2	143
-8	624698

Waves: all

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

#### pgmode – Interview Method

---

100	[100] With Interviewer Assistance	10869
110	[110] Oral Interview	159274
120	[120] Written Ques. Interviewer	128312
130	[130] Mixed Type	0
131	[131] Written Ques. No Interviewer	24592
132	[132] Oral And Written	18901
133	[133] Proxy	155
134	[134] Third Person Present	3227
135	[135] No Third Person Present	20991
140	[140] CAPI - Since 1998 (O)	216975
150	[150] Cawi Since 2014 (BE)	7107
200	[200] Telephone Assistance	0
210	[210] Written, By Mail	66428



220	[220] Telephone Interview	247
-1	[-1] No Answer	866
-2	[-2] Does not apply	0
-3	[-3] Answer improbable	19
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-8	[-8] Question this year not part of Survey program	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
bpgen	11090
cpgen	10646
dpgen	10516
epgen	10023
fpgen	9710
gpgen	13972
... (4 rows omitted)	53662
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