

SOEP Survey Papers

Series D – Variable Descriptions and Coding

SOEP – The German Socio-Economic Panel Study at DIW Berlin

2014

SOEP 2013 – Documentation of Person-related Status and Generated Variables in PGEN for SOEP v30

SOEP Group

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

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

The SOEP Survey Papers is comprised of the following series:

Series A – Survey Instruments (Erhebungsinstrumente)

Series B – Survey Reports (Methodenberichte)

Series C – Data Documentations (Datendokumentationen)

Series D – Variable Descriptions and Coding

Series E – SOEPmonitors

Series F – SOEP Newsletters

Series G – General Issues and Teaching Materials

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

Editors:

Prof. Dr. Gert G. Wagner, DIW Berlin and Technische Universität Berlin

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

Please cite this paper as follows:

SOEP Group. 2014. SOEP 2013 – Documentation of Person-related Status and Generated Variables in PGEN for SOEP v30. SOEP Survey Papers 250: Series D. Berlin: DIW/SOEP

ISSN: 2193-5580 (online)

Contact: DIW Berlin
SOEP
Mohrenstr. 58
10117 Berlin

Email: soepapers@diw.de

Jürgen Schupp, Martin Kroh, Jan Goebel, Carsten Schröder, Elisabeth Bügelmayer, Marco Giesselmann, Markus Grabka, Peter Krause, Simon Kühne, Elisabeth Liebau, David Richter, Rainer Siegers, Paul Schmelzer, Christian Schmitt, Daniel Schnitzlein, Ingrid Tucci, Knut Wenzig
(SOEP Group)

SOEP 2013 – DOCUMENTATION OF PERSON-RELATED STATUS AND GENERATED VARIABLES IN PGEN FOR SOEP v30

Berlin, 2014

DOI: 10.5684/soep.v30

<http://www.diw.de/doi/soep.v30>

List of variables :

ERWTYP\$\$--Type of occupation.....	3
JOBCH\$\$-- Job change.....	4
EMPLST\$\$ -- Employment status.....	5
LFSS\$ -- Labor force status.....	6
EXPFT\$\$-- Working experience, full-time employment.....	7
EXPPT\$\$ -- Working experience, part-time employment.....	8
EXPUES\$ -- Unemployment experience.....	9
\$ERWZEIT -- Length of time with firm.....	10
ERLJOB\$\$ -- Working in occupation trained for.....	11
AUSB\$\$ -- Required job training.....	12
BETR\$\$ -- Size of company.....	14
ALLBET\$\$ -- Core size category of the company.....	16
OEFFD\$\$ -- Civil service.....	17
STIB\$\$ -- Occupational position.....	18
NACE\$\$ -- Two-digit NACE Industry - Sector.....	20
AUTONOS\$ -- Autonomy in occupational activity.....	22
\$VEBZEIT -- Agreed weekly working hours.....	23
\$STATZEIT -- Actual weekly working hours.....	24
\$UEBSTD -- Overtime per week.....	25
ISEI\$\$ -- International Socio-Economic Index of Occupational Status.....	29
EGP\$\$ -- Erikson and Goldthorpe Class Category.....	30
SIOPSS\$ -- Treimans Standard Int. Occupation Prestige Score.....	32
MPSS\$ -- Magnitude-Prestige Scala - Wegener.....	33
PARTZ\$\$ -- Partner indicator.....	34
PARTNR\$\$ -- Person ID number of partner.....	35
\$FAMSTD -- Marital status in survey year.....	36
NATION\$\$ -- Citizenship - nationality.....	37
\$PSBIL -- Diplomas/degrees from secondary/tertiary education.....	38
\$PSBILO -- Secondary school degree/diploma East Germany.....	44
\$PSBILA -- Secondary school degrees/diplomas abroad.....	45
\$PBBILO1 -- Vocational degree attained.....	46
\$PBBILO -- Vocational degree attained - East.....	47
\$PBBILO2 -- Completed college education.....	48
\$PBBILO3 -- No vocational degree.....	49
\$PBBILA -- Vocational degree abroad.....	50
\$BILZEIT -- Amount of education or training (in years).....	51
\$ISCED -- Highest degree/diploma attained, ISCED-1997 classification.....	53
\$CASMIN -- Highest degree/diploma according to CASMIN classification.....	55
MONTH\$\$ -- Month of interview.....	57
MODE\$\$ -- Interview method.....	58
LABGRO\$\$ -- Current gross labor income in euros (generated).....	59
IMPGRO\$\$ -- Imputation flag for LABGRO\$\$.....	60
LABNET\$\$ -- Current net labor income (generated) in euros.....	61
IMPNET\$\$ -- Imputation flag for current net labor income (LABNET\$\$).....	62
SNDJOB\$\$ -- Current gross labor income in euros from second job (generated)..	63
IMPSND\$\$ -- Imputation flag for SNDJOB\$\$.....	64
FIELD\$\$ -- Field of tertiary education.....	65
DEGREE\$\$ -- Type of tertiary degree.....	68
TRAINA\$\$ -- Apprenticeship - two-digit occupation KldB92.....	70
TRAINB\$\$ -- Vocational school - two-digit occupation KldB92.....	72
TRINC\$\$ -- Higher vocational school - two-digit occupation KldB92.....	72
TRAIND\$\$ --Civil servant training - two-digit occupation KldB92.....	72
FDT_F\$\$ -- Data source FIELD, DEGREE, TRAIN.....	73
Variables no longer distributed:.....	76
ISCOU\$\$.....	76
ISCO\$\$.....	76
ISCOH\$\$.....	76
BRANCH\$\$.....	76

ERWTYP\$\$--Type of occupation

Var Label : ERWTYP\$\$ ""
Value Label : ERWTYP\$\$ (-1)"no answer"
(1)"not employed"
(2)"not employed (first-time respondent)"
(3)"employed (first-time respondent)"
(4)"employed, no change"
(5)"employed, no info if change"
(6)"employed, with change or first time employed"
(7)"employed, in part-time work with approaching retirement"

Var format : ERWTYP\$\$ (I2)
19\$\$ - Year : \$\$=84..13

Comment: This variable is generated from the question on whether a respondent has changed jobs since the beginning of the previous year, which is a central filter variable in the questionnaire. In years with a partial survey – 1985, 1986, 1987, 1988, 1990 (West), 1992 (West), 1994, 1996, 1999, 2001, 2003, 2005, 2006, 2008, 2010 and 2012 – only employed persons who changed jobs and first-time respondents are asked to provide up-to-date information on time-invariant job characteristics. Therefore, in years with a partial survey, for persons without a job change, many of the generated variables related to the job contain information from the previous year. The respondents from the newly introduced migration sample were not asked whether they have changed jobs since the beginning of the previous year, so only Codes (1), (5) and (7) were assigned.

The variable ERWTYP\$\$ was originally created to integrate the "blue" (first-time respondent) and "green" (follow-up respondent) questionnaires used up to 1993 to differentiate between employed persons *with* and *without* a change of job (as a central filter variable). Since the 1994 wave, there has been only one questionnaire for all respondents – both first-time and follow-up respondents. For this reason, since the 1994 wave, Codes (2) and (3) are no longer assigned. Codes (1), (4), and (5) have been assigned since 1994 to first-time respondents (who would have previously received the blue questionnaire). The variable was recalculated for all waves and is assigned the code (6) only when the respondent started a new job since the last interview or was hired for his or her first job ever. The variable also includes a new code (7) since wave T (2003) for employed persons in a phased retirement scheme (*Altersteilzeit*) whose current actual working hours are zero.

An alternative variable is JOBCH\$\$ (see below), which is an improved version of ERWTYP\$\$, as it is generated in a longitudinally consistent way and contains an additional category for first-time employed persons.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

JOBCH\$\$-- Job change

Var Label : JOBCH\$\$ **"Job change"**
Value Label : JOBCH\$\$ (-1)"no answer"
(1)"not employed"
(2)"employed, no change"
(3)"employed, no info if change"
(4)"employed, with change"
(5)"first time employed"

Var format : JOCHCH\$\$ (I2)
19\$\$ - Year : \$\$=84..13

Comment: 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.

JOCHCH\$\$ is generated based on the central filter variable, which indicates whether a respondent has changed jobs since the beginning of the previous year. Information on the date of job change is then combined with interview month of the previous year's interview to identify whether a new job change has taken place since the previous interview.

Hence, JOBCH\$\$ is a modified version of the variable ERWTYP\$\$ which indicates whether a respondent has changed jobs since the beginning of the previous year. Unlike ERWTYP\$\$, the variable is calculated for all waves, and the codes are assigned independently of the respondent being a first-time or follow-up respondent.

In addition to ERWTYP\$\$, 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.

In addition to ERWTYP\$\$, the variable is designed to provide consistent longitudinal information on job changes. The JOBCH\$\$ variable is generated by correcting the original job change information in various ways:

1. We check whether the job changes stated by a respondent in two 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".

The respondents from the newly introduced migration sample

were not asked whether they have changed jobs since the beginning of the previous year, therefore the generation of JOBCH\$\$ was modified:

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).

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

EMPLST\$\$ -- Employment status

Var Label : EEMPLST\$\$ "Employment status"
Value Labels : EEMPLST\$\$ (-3)"implausible answer"
(-1)"no answer"
(1)"full-time employment"
(2)"regular part-time employment"
(3)"vocational training"
(4)"marginal, irregular part-time employment"
(5)"not employed"
Var format : EEMPLST\$\$ (I2)
19\$\$ - year : \$\$=84..97

Var Label : EEMPLST\$\$ "Employment status"
Value Labels : EEMPLST\$\$ (-3)"implausible answer"
(-1)"no answer"
(1)"full-time employment"
(2)"regular part-time employment"
(3)"vocational training"
(4)"marginal, irregular part-time employment"
(5)"not employed"
(6)"sheltered workshop"
Var format : EEMPLST\$\$ (I2)
19\$\$ - year : \$\$=98..13

Comment: This variable is generated from the annual question on current employment status, which has a central filter function in the questionnaire to separate employed people from non-employed people for further questions. It is designed to provide consistent longitudinal data on employment status across all waves. Since the beginning of the SOEP in the year 1984, a consistent status variable has been used to differentiate among different types of employment status. The category "not employed" comprises non-working individuals, those in military/community service, those on maternity leave, and employed persons in a phased retirement scheme (*Altersteilzeit*) whose current actual working hours are zero. From 1998 on, the additional category "sheltered workshop" is included for disabled persons in sheltered employment.

EMPLST\$\$ supplements the variable LFSS\$\$, which differentiates among persons who are not employed.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

LFSS\$ -- Labor force status

Var Label : LFSS\$ "Labor force status"
Value Label : LFSS\$ (-1)"no answer"
(1)"non-working without further information"
(2)"non-working, and older than 65"
(3)"non-working, in training program"
(4)"non-working, on maternity leave"
(5)"non-working, in military/community service"
(6)"non-working, and registered unemployed"
(8)"non-working, but sometimes second job"
(9)"non-working, but working past 7 days"
(10)"non-working, but regular second job"
(11)"working"
(12)"working, but non-working past 7 days"

Var format : LFSS\$ (I2)
19\$\$ - Year : \$\$=84..13

Comment: 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.

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

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

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

LFSS\$ supplements the variable EMPLST\$, which differentiates among persons who are employed. The respondents from the newly introduced migration sample were not asked about secondary jobs, so Codes (8) and (10) were not assigned to them.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

EXPFT\$\$ -- Working experience, full-time employment

Var Label : EXPFT\$\$ "Working experience, full-time employment"
Value Labels : EXPFT\$\$ (-3)"implausible answer"
(-1)"no answer"

Var format : EXPFT\$\$ (I3)
19\$\$ - year : \$\$=84..13

Comment: Full-time working experience:
This variable reflects the total length of full-time employment in the respondent's career up to the point of the interview in a given year. The variable is created by combining monthly information on employment status from the calendar dataset *ARTKALEN* (which provides monthly information on activity status since an individual entered the SOEP) and annual information from the biographical dataset *PBIOSPE* (which provides information on activity status over the individual's life course). EXPFT\$\$ gives the length of time in years with months in decimal form.

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

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

Please also see EXPPT\$\$ and EXPUE\$\$.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

EXPPT\$\$ -- Working experience, part-time employment

Var Label : EXPPT\$\$ "Working experience, part-time employment"
Value Labels : EXPPT\$\$ (-3)"implausible answer"
(-1)"no answer"

Var format : EXPPT\$\$ (I3)
19\$\$ - year : \$\$=84..13

Comment: Part-time working experience:
This variable reflects the total length of part-time employment in the respondent's career up to the point of the interview in a given year. The variable is created by combining monthly information on employment status from the calendar dataset *ARTKALEN* (which provides monthly information on activity status since an individual entered the SOEP) and annual information from the biographical dataset *PBIOSPE* (which provides information on activity status over the life course of an individual). EXPPT\$\$ gives the length of time in years with months in decimal form.

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

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

Please also see EXPFT\$\$ and EXPUE\$\$.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

EXPUE\$\$ -- Unemployment experience

Var Label : EXPUE\$\$ "Unemployment experience"
Value Labels : EXPUE\$\$ (-3)"implausible answer"
(-1)"no answer"

Var format : EXPUE\$\$ (I3)
19\$\$ - year : \$\$=84..13

Comment: Unemployment experience:
This variable reflects the total length of unemployment in the respondent's career up to the point of the interview in a given year. The variable is created by combining monthly information on employment status from the calendar dataset *ARTKALEN* (which provides monthly information on activity status since an individual entered the SOEP) and annual information from the biographical dataset *PBIOSPE* (which provides information on activity status over the life course of an individual). EXPUE\$\$ gives the length of time in years with months in decimal form.

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

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

Please also see EXPFT\$\$ and EXPPT\$\$.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

\$ERWZEIT -- Length of time with firm

Var Label : \$ERWZEIT **"Length of time with firm"**
Value Labels : \$ERWZEIT (-3)"implausible answer"
(-2)"does not apply - not employed"
(-1)"no answer"

Var format : \$ERWZEIT (I3)
19\$\$ - Jahr : \$\$=84..13

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

The variable provides consistent longitudinal information on the length of time with the same employer. Data that show longitudinal inconsistencies are corrected.

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: SOEPmail (Tel. +49-30-89789-292)

ERLJOB\$\$ -- Working in occupation trained for

Var Label : ERLJOB\$\$ **"Working in occupation trained for"**
Value Label : ERLJOB\$\$ (-2)"does not apply - not employed"
(-1)"no answer"
(1)"yes"
(2)"no"
(3)"currently in training"
(4)"has no job training"

Var format : ERLJOB\$\$ (I2)

19\$\$ - year : \$\$=84..13

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

Not all employed persons are asked this question 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 and 2012 – 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, ERLJOB\$\$ usually contains available information from the previous year for persons without a job change. For some persons without a job change who updated the information on their current occupation without being asked, up-to-date information is available.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

AUSB\$\$ -- Required job training

Var Label : AUSB\$\$ **"Required job training"**
Value Label : AUSB\$\$ (-2)"does not apply - not employed"
(-1)"no answer"
(1)"no training required"
(2)"brief on-the-job training"
(3)"extensive on-the-job training"
(4)"attended courses"
(5)"completed vocational training"
(7)"Fachhochschule or university degree"
Var format : AUSB\$\$ (I2)
19\$\$ - Year : \$\$=84..89

Var Label : AUSB\$\$ **"Required job training"**
Value Label : AUSB\$\$ (-2)"does not apply - not employed"
(-1)"no answer"
(1)"no training required"
(2)"brief on-the-job training"
(3)"extensive on-the-job training"
(4)"attended courses"
(5)"completed vocational training"
(6)"completed technical school (East), 1990-1993"
(7)"Fachhochschule or university degree"
Var format : AUSB\$\$ (I2)
19\$\$ - Year : \$\$=90..98

Var Label : AUSB\$\$ **"Required job training"**
Value Label : AUSB\$\$ (-2)"does not apply - not employed"
(-1)"no answer"
(1)"no training required"
(2)"brief on-the-job training"
(3)"extensive on-the-job training"
(4)"attended courses"
(5)"completed vocational training"
(8)"Fachhochschule degree"
(9)"University degree"
Var format : AUSB\$\$ (I2)
19\$\$ - Jahr : \$\$=99..13

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

Not all employed persons are asked this question 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 and 2012 - 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, AUSB\$\$ usually contains available information from the previous year for persons without a job change. For some respondents without a job change who updated the information on their current occupation without being asked, up-to-date information is available.

The answer option "completed technical school" was only used from 1990 to 1993 in the East German version of the questionnaire. Since not all employed people are asked the question about required training every year, the value (6) of the variable AUSB\$\$ is valid up to 1996.

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

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

BETR\$\$ -- Size of company

Var Label : BETR\$\$ **"Size of company"**
Value Label : BETR\$\$ (-2)"does not apply - not employed"
(-1)"no answer"
(4)"less than 20"
(8)"20 to 200"
(9)"200 to 2000"
(10)"2000 or more"
(11)"Self-employed without other employees"
Var format : BETR\$\$ (I2)
19\$\$ - Year : \$\$=84..90

Var Label : BETR\$\$ **"Size of company"**
Value Label : BETR\$\$ (-2)"does not apply - not employed"
(-1)"no answer"
(1)"less than 5"
(4)"up to 1990: less than 20"
(5)"5 to 20"
(8)"20 to 200"
(9)"200 to 2000"
(10)"2000 or more"
(11)"Self-employed without other employees"
Var format : BETR\$\$ (I2)
+19\$\$ - Year : \$\$=91..98

Var Label : BETR\$\$ **"Size of company"**
Value Label : BETR\$\$ (-2)"does not apply - not employed"
(-1)"no answer"
(1)"less than 5"
(4)"up to 1990: less than 20"
(5)"5 to 20"
(6)"20 to 100"
(7)"100 to 200"
(8)"up to 1998: 20 to 200"
(9)"200 to 2000"
(10)"2000 or more"
(11)"Self-employed without other employees"
Var format : BETR\$\$ (I2)
+19\$\$ - Year : \$\$=99..04

Var Label : BETR\$\$ **"Size of company"**
Value Label : BETR\$\$ (-2)"does not apply - not employed"
(-1)"no answer"
(1)"less than 5"
(2)"5 to 10"
(3)"11 to 20"
(4)"up to 1990: less than 20"
(5)"1991-2004: 5 to 20"
(6)"20 to 100"
(7)"100 to 200"
(8)"up to 1998: 20 to 200"
(9)"200 to 2000"
(10)"2000 or more"
(11)"Self-employed without other employees"
Var format : BETR\$\$ (I2)
+20\$\$ - Year : \$\$=05..13

Comment: This variable is designed to offer annual data on company size for all employed persons.

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

Not all employed persons are asked this question on an annual basis. In years with a partial survey - in 1985, 1986, 1987, 1988, 1990 (West), 1992 (West), 1994, 1996, 1999, 2001, 2003, 2005, 2006, 2008, and 2010 (but not in 2012) - 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, BETR\$\$ usually contains available information from the previous year for persons without a job change. For some persons without a job change who updated the information on their current occupation without being asked, up-to-date information is available.

In 2012, the questionnaire provides a one-time-only information on the size of the local establishment in addition to the size of the entire company (BETR\$\$). The enriched questionnaire revealed that in previous interviews, persons have mistakenly provided information on the local establishment size instead of the entire company size, especially if their entire company had 2000 and more employees. Due to the importance of longitudinal consistency, these persons 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 ALLBET\$\$ for a broader categorization of the firm size, which is appropriate for analyses that include all sample years.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

ALLBET\$\$ -- Core size category of the company

Var Label : ALLBET\$\$ **"Core size category of the company"**
Value Label : ALLBET\$\$ (-2)"does not apply - not employed"
(-1)"no answer"
(1)"fewer than 20"
(2)"20 to 200"
(3)"200 to 2000"
(4)"2000 or more"
(5)"Self-employed without other employees"

Var format : ALLBET\$\$ (I2)
19\$\$ - Year : \$\$=84..13

Comment: This variable is designed to provide annual data on the core size category of the company for all employed persons. Since respondents were asked about company size in more detailed form in the years 1991, 1999 and 2005 (see also BETR\$\$), the variable ALLBET\$\$ contains the lowest common denominator of the variable BETR\$\$, i.e., the firm size categories available across all SOEP waves. This broader categorization corresponds to the values of variables BETR84 to BETR90 and offers a variable that is consistent across all waves.

Not all employed persons are asked the question on firm size on an annual basis. In years with a partial survey - in 1985, 1986, 1987, 1988, 1990 (West), 1992 (West), 1994, 1996, 1999, 2001, 2003, 2005, 2006, 2008, and 2010 (but not in 2012)- 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, ALLBET\$\$ usually contains available information from the previous year for persons without a job change. For some persons without a job change who updated the information on their current occupation without being asked, up-to-date information is available.

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

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

OEFFD\$\$ -- Civil service

Var Label : OEFFD\$\$ **"Civil service"**
Value Label : OEFFD\$\$ (-2)"does not apply - not employed"
(-1)"no answer"
(1)"yes"
(2)"no"

Var format : OEFFD\$\$ (I2)
19\$\$ - Year : \$\$=84..13

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

Not all employed persons are asked this question 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 and 2012 - 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, OEFFD\$\$ usually contains available information from the previous year for persons without a job change. For some persons without a job change who updated the information on their current occupation without being asked, up-to-date information is available.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

STIB\$\$ -- Occupational position

Var Label	:	stib\$\$	"Occupational position"
Value Labels	:	stib\$\$	(1-640)(see below)
Var format	:	STIB\$\$	(I3)
19\$\$ - Year	:	\$\$=84..13	

(-2) Does not apply		
(-1) No Answer	(422)	Free-Lance Professional, LE 9 Employees
(10) Not Employed	(423)	Free-Lance Professional, GT 9 Employees
(11) In Education	(430)	Other Self-Employed No Or LE 9 Employees
(12) Unemployed, Not Employer	(431)	Other Self-Employed No Employees
(13) Pensioner	(432)	Other Self-Employed LE 9 Employees
(15) Military, Community Service	(433)	Other Self-Employed GT 9 Employees
(110) Apprentice	(440)	Help In Family Business
(120) Apprentice, Trainee Industry Technology	(510)	Foreman
(130) Apprentice, Trainee Trade and Commerce	(520)	Employee with Simple Tasks
(140) Trainee, Intern	(520)	W-Collar Worker with Simple Tasks
(150) Research assistant	(521)	Untrained Employee with Simple Tasks
(210) Untrained Worker	(521)	Untrained W-Collar Worker with Simple Tasks
(220) Semi-Trained Worker	(522)	Trained Employee with Simple Tasks
(230) Trained Worker	(522)	Trained W-Collar Worker with Simple Tasks
(240) Foreman, Team Leader	(530)	Qualified Professional
(250) Foreman	(540)	H. Qualified Professional
(310) Agricultural Worker	(550)	Managerial
(320) Agricultural Specialist	(610)	Low-Level Civil Service
(330) Agricultural Foreman	(620)	Middle-Level Civil Service
(340) Agricultural Manager	(630)	High-Level Civil Service
(410) Self-Employed Farmer	(640)	Executive Civil Service
(411) Self-Employed Farmer, No Employees		
(412) Self-Employed Farmer LE 9 Employees		
(413) Self-Employed Farmer GT 9 Employees		
(420) Free-Lance Professional		
(421) Free-Lance Professional, No Employees		

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

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

The categories (150) and (310) to (340) were only assigned to respondents in East Germany in 1990. In STIB\$\$, non-working persons are only assigned to the category (13) "pensioner" if they are recipients of retirement pension or if they are recipients of widow's pension AND are older than 60 years. Moreover, if there is missing information on pension receipt, additional information from *ARTKALEN* (retrospective information

from the activity calendar for the previous year) is used in the generation process to determine if a person was in retirement or early retirement (*Vorruhestand*) at the time of the interview.

The code (-2) is assigned to first time respondents aged 16 or 17 who answer the youth questionnaire (since wave W (2006)).

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

NACE\$\$ -- Two-digit NACE Industry - Sector

Var Label : NACE\$\$ "Two-digit NACE Industry - Sector"
 Value Label : NACE\$\$ (1-100) (see below)

Var format : NACE\$\$ (I2)
 19\$\$ - Year : \$\$=84..13

- (-2) Does not apply - not employed
- (-1) No Answer
- (1) Agriculture, Hunting, Related Service Activities
- (2) Forestry, Logging, Related Service activities
- (5) Fishing, Operation Of Fish Hatcheries And Fish Farms
- (10) Mining Of Coal And Lignite; Extraction Of Peat
- (11) Extraction Of Crude Petroleum And Natural Gas
- (12) Mining Of Uranium And Thorium Ores
- (13) Mining Of Metal Ores
- (14) Other Mining And Quarrying
- (15) Manuf. Food Products And Beverages
- (16) Manuf. Tobacco Products
- (17) Manuf. Textiles
- (18) Manuf. Wearing Apparel; Dressing And Dyeing Of Fur
- (19) Tanning, Dressing Of Leather; Manuf. luggage, Footwear
- (20) Manuf. Wood Products, Except Furniture
- (21) Manuf. Pulp, Paper And Paper Products
- (22) Publishing, Printing And Reproduction Of Recorded Media
- (23) Manuf. Coke, Refined Petroleum Prod, Nuclear Fuel
- (24) Manuf. Chemicals And Chemical Products
- (25) Manuf. Rubber And Plastic Products
- (26) Manuf. Other Non-metallic Mineral Products
- (27) Manuf. Basic Metals
- (28) Manuf. Fabricated Metal Prod., Ex. Machinery And Equip
- (29) Manuf. Machinery And Equipment NEC
- (30) Manuf. Office Machinery And Computers
- (31) Manuf. Electrical Machinery And Apparatus NEC
- (32) Manuf. Radio, Television And Communication Equipment
- (33) Manuf. Medical, Precision And Optical Instruments
- (34) Manuf. Motor Vehicles, Trailers And Semi-trailers
- (35) Manuf. Other Transport Equipment
- (36) Manuf. Furniture; Manufacturing NEC
- (37) Recycling
- (40) Electricity, Gas, Steam And Hot Water Supply
- (41) Collection, Purification And Distribution Of Water
- (41) Collection, Purification And Distribution Of Water
- (45) Construction
- (50) Sale, Maint., Repair Motor Vehicles; Retail Car Gas
- (51) Wholesale Trade, Commission Trade, Ex. Motor Vehicles
- (52) Retail, Ex. Motor vehicles, Motorcycles; Repair
- (55) Hotels And Restaurants
- (60) Land Transport; Transport Via Pipelines
- (61) Water Transport
- (62) Air Transport
- (63) Supporting, Aux. Transport Activities; Travel agencies
- (64) Post And Telecommunications
- (65) Financial Intermediation, Ex. Insurance, Pension Funding
- (66) Insurance And Pension Funding, Ex. Compulsory SocSec
- (67) Activities Auxiliary To Financial Intermediation
- (70) Real Estate, Property Activities
- (71) Renting Of Machinery, Equip Wo. Oper., Pers,HH Goods
- (72) Computer And Related Activities
- (73) Research And Development
- (74) Other Business Activities
- (75) Public Administration And Defense; Compulsory SocSec
- (80) Education
- (85) Health And Social Work
- (90) Sewage And Refuse Disposal, Sanitation And Related
- (91) Activities Of Membership Organizations NEC.
- (92) Recreational, Cultural And Sporting Activities
- (93) Other Service Activities
- (95) Private Households With Employed Persons
- (96) Industry - NEC
- (97) Handcraft, Trade - NEC
- (98) Services - NEC
- (99) Extra-territorial Organizations And Bodies
- (100) Manufacturing - NEC

Comment: This variable is designed to provide annual data on the industry of economic activity for all employed persons according to the *Statistical Classification of Economic Activities in the European Community (Nomenclature des statistiques des activités économiques de la Communauté européenne - NACE)*. Respondents answer the question in their own words regarding the industry in which they are currently working, and this response is entered into a blank in the questionnaire.

In order to facilitate international comparability, the European industry standard classification system is used by Infratest Sozialforschung to recode this information. This recoding has been documented in Hartmann/Schütz 2002.

The codes in *NACE Rev.1* also correspond to *ISIC Rev.3* (*International Standard Classification of All Economic Activities*). With the 2001 data distribution, the sector codes formerly used in the SOEP were completely recoded to the *NACE* classification. Please note that special codes 96-98 as well as 100 were assigned by Infratest Sozialforschung whenever respondents did not provide a more detailed answer.

Not all employed persons are asked this question 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 and 2012 - 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, *NACE* usually contains available information from the previous year for persons without a job change. For some persons without a job change who updated the information on their current occupation without being asked, up-to-date information is available.

The respondents from the newly introduced migration sample were not asked in which industry they are currently working, so (-1) was assigned to all employed respondents from the migration sample.

Detailed description 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: SOEPmail (Tel. +49-30-89789-292)

AUTONO\$\$ -- Autonomy in occupational activity

Var Label : AUTONO\$\$ **"Autonomy in occupational activity"**
Value Labels : AUTONO\$\$ (-2)"does not apply - not employed"
(-1)"no answer"
(0)"apprentice, intern, unpaid trainee"
(1)"low autonomy"
(2)"low-medium autonomy"
(3)"medium autonomy"
(4)"medium-high autonomy"
(5)"high autonomy"

Var format : AUTONO\$\$ (I2)
19\$\$ - Year : \$\$=84..13

Comment: This variable gives the occupational autonomy for all employed persons. It offers an alternative to the ISCO-based scales on occupational status (ISEI\$\$), class (EGP\$\$), or prestige (SIOP\$\$). AUTONO\$\$ is the simplest variable based on the scales of "occupational position" in terms of its construction, and strongly correlated with the Treiman Prestige Scale (SIOP\$\$).

The basis for the "autonomy in occupational activity" scale is the classification of occupational position. Self-employed persons are categorized according to the size of the company (with the exception of farmers, who are all classified within the same category of autonomy, independent of farm size in hectares). Civil servants are differentiated according to the civil service laws defining each kind of activity and the amount of autonomy connected to it. Workers are differentiated according to their vocational training, and thus categorized hierarchically according to the different tasks they can be expected to carry out and the different amounts of responsibility associated with each task. Similarly, salaried employees are classified according to how differentiated their tasks are and how much responsibility is associated with each.

The value "1" is assigned mainly to manual workers with a low level of status and a low level of autonomy. Group 2 encompasses work in production, services demanding a minimal level of specialization, and farm work. Activities that require completion of the middle track of secondary education and entail a limited amount of responsibility are classified in Group 3. Group 4 includes activities carried out either with or without supervision that require a degree from a college of applied sciences or university, but are not very high in prestige. Managers and freelance academics are both placed in Group 5 (highest autonomy). Depending on the number of employees, self-employed are categorized in Group 3, Group 4, or Group 5.

Detailed description: Hoffmeyer-Zlotnik, Jürgen H.P., and Alfons J. Geis (2003) Berufs-
klassifikation und Messung des beruflichen Status/ Prestige. In: ZUMA-
Nachrichten 52, Jg. 27, Mai 2003. pp. 125-138.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

\$VEBZEIT -- Agreed weekly working hours

Var Label : \$VEBZEIT **"Agreed weekly working hours"**
Value Labels : \$VEBZEIT (-3)"implausible answer"
(-2)"does not apply - not employed, self-
employed, no set hours"
(-1)"no answer"

Var format : \$VEBZEIT (I3)
19\$\$ - Year : \$\$=84..13

Comment: This variable is designed to offer annual data on agreed weekly working hours. The variable takes into account only those persons who were in dependent employment (not self-employed) at the time of the survey. Agreed weekly working hours were asked up to 1989 only in full hours, and from 1990 on in three-digit form (counting the first digit after the decimal point).

The value (-2) is assigned to non-employed people, employees without set hours and to self-employed people, including self-employed farmers, freelancers, and other self-employed persons. 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 \$STATZEIT and \$UEBSTD.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

\$STATZEIT -- Actual weekly working hours

Var Label : \$STATZEIT "Actual weekly working hours"
Value Labels : \$STATZEIT (-3)"implausible answer"
(-2)"does not apply - not employed"
(-1)"no answer"

Var format : \$STATZEIT (I3)
\$\$ - Year : \$\$=84..13

Comment: This variable is designed to offer annual data on actual weekly working hours (including overtime) for all persons employed at the time of the survey (including the self-employed). The data are obtained by asking respondents how many hours they work *on average* per week. Actual weekly working hours were asked up to 1989 only in full hours, and from 1990 on as a three-digit number (counting the first digit after the decimal point).

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

Please also see \$VEBZEIT and \$UEBSTD.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

\$UEBSTD -- Overtime per week

Var Label : \$UEBSTD "Overtime per week"
Value Labels : \$UEBSTD (-3)"implausible answer"
(-2)"does not apply - not employed, self-
employed, no set hours"
(-1)"no answer"

Var format : \$UEBSTD (I3)
\$\$ - Year : \$\$=84..13

Comment: This variable is designed to offer annual data on overtime per week for all persons in dependent employment at the time of the survey. The data is obtained by asking respondents how many overtime hours they worked *in the month before the survey*. The number of monthly overtime hours is then converted into weekly overtime by dividing the number given by 4.3. Since \$UEBSTD refers to weekly overtime during the *last month*, the number may deviate from the difference between average actual weekly working hours and the agreed weekly working hours.

In the years 1984, 1985 and 1987, respondents were not asked about number of hours of overtime per week. The variables were therefore generated using the difference between average actual weekly working hours and agreed weekly working hours. Respondents were asked for the number of overtime hours up to 1989 only in full hours, and from 1990 on, as a three-digit number (counting the first digit after the decimal point).

The value (-2) is assigned to non-employed people, employees without set hours and to self-employed people, including self-employed farmers, freelancers, and other self-employed persons. If persons helping out in family businesses report overtime hours, we assign a non-missing value. For implausible answers (agreed-upon weekly working time or actual weekly working time of more than 80 hours per week AND weekly overtime of more than 10 hours we assign the value (-3).

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

Please also see \$VEBZEIT and \$TATZEIT.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

JOBEND\$\$ -- Reasons for occupational change

Var Label : JOBEND\$\$ **"Reasons for occupational change"**
 Value Label : JOBEND\$\$ (-2)"does not apply - not employed"
 (-1)"no answer"
 (1)"Terminated by employer"
 (2)"Temporary contract expired"
 (3)"Education, training completed"
 (4)"Own resignation"
 (5)"Mutual termination"
 (6)"Employee requested transfer"
 (7)"Company transferred employee"
 (8)"Ended self-employment"
 (15)"Other incl. early retirement, company
 closed, old-age pension, leave of absence/
 sabbatical(1985-1986)"

Var format : JOBEND\$\$
 19\$\$ - Year : \$\$=85,86

Var Label : JOBEND\$\$ **"Reasons for occupational change"**
 Value Label : JOBEND\$\$ (-2)"does not apply - not employed"
 (-1)"no answer"
 (1)"Terminated by employer"
 (2)"Temporary contract expired"
 (3)"Education, training completed"
 (4)"Own resignation"
 (5)"Mutual termination"
 (6)"Employee requested transfer"
 (7)"Company transferred employee"
 (8)"Ended self-employment"
 (10)"Took early retirement"
 (16)"Other incl. company closed, old-age pension,
 leave of absence/sabbatical (1987-1990)"

Var format : JOBEND\$\$
 +19\$\$ - Year : \$\$=87..90

Var Label : JOBEND\$\$ **"Reasons for occupational change"**
 Value Label : JOBEND\$\$ (-2)"does not apply - not employed"
 (-1)"no answer"
 (1)"Terminated by employer"
 (2)"Temporary contract expired"
 (3)"Education, training completed"
 (4)"Own resignation"
 (6)"Employee requested transfer"
 (7)"Company transferred employee"
 (8)"Ended self-employment"
 (10)"Took early retirement"
 (11)"Company closed down"
 (12)"Old-age pension"
 (14)"Leave, maternity leave and parental leave
 (1991-1998), since 2011"
 (17)"Other incl. mutual termination (1991-1998)"

Var format : JOBEND\$\$
 +19\$\$ - Year : \$\$=91..98

Var Label : JOBEND\$\$ **"Reasons for occupational change"**
 Value Label : JOBEND\$\$ (-2)"does not apply - not employed"
 (-1)"no answer"
 (1)"Terminated by employer"
 (4)"Own resignation"
 (5)"Mutual termination"
 (8)"Ended self-employment"
 (9)"Temporary contract expired or education/
 training completed"
 (12)"Old-age pension"
 (13)"Leave of absence/sabbatical (1999-2010)"

Var format : JOBEND\$\$
 +19\$\$ - Year : \$\$=99,00

Var Label : JOBEND\$\$ "Reasons for occupational change"
 Value Label : JOBEND\$\$ (-2)"does not apply - not employed"
 (-1)"no answer"
 (1)"Terminated by employer"
 (4)"Own resignation"
 (5)"Mutual termination"
 (8)"Ended self-employment"
 (9)"Temporary contract expired or education/
 training completed"
 (11)"Company closed down"
 (12)"Old-age pension"
 (13)"Leave of absence/sabbatical (1999-2010)"
 Var format : JOBEND\$\$
 +20\$\$ - Year : \$\$=00..10

Var Label : JOBEND\$\$ "Reasons for occupational change"
 Value Label : JOBEND\$\$ (-2)"does not apply - not employed"
 (-1)"no answer"
 (1)"Terminated by employer"
 (4)"Own resignation"
 (5)"Mutual termination"
 (8)"Ended self-employment"
 (9)"Temporary contract expired or education/
 training completed"
 (11)"Company closed down"
 (12)"Old-age pension"
 (14)"Leave, maternity leave and parental leave
 (1991-1998), since 2011"
 Var format : JOBEND\$\$
 +20\$\$ - Year : \$\$=11..13

Comment: 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 1985-1991 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.

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.

The respondents from the newly introduced migration sample were not asked about an occupational change, so no values were assigned to them.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

COREJOBEND\$\$ -- Core reasons for occupational change

Var Label : COREJOBEND\$\$ "Core reasons for occupational change"
 Value Label : COREJOBEND\$\$ (-2)"does not apply - not employed"
 (-1)"no answer"
 (1)"Terminated by employer"
 (2)"Temporary contract expired or education/
 training completed"
 (3)"Own resignation"
 (4)"Ended self-employment"
 (11)"Other"
 Var format : COREJOBEND\$\$
 19\$\$ - Year : \$\$=85..13

Comment: This variable is designed to offer annual data on core reasons for an occupational change for all formerly employed persons, persons with a job change or persons on leave. 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 year to the following. Due to the fact that in any years respondents were asked about reasons for change with more or less given answers (see also JOBEND\$\$), the variable COREJOBEND\$\$ contains the lowest common denominator, i.e. values of the variable JOBEND\$\$, available across all SOEP waves. This broader categorization offers a variable that is consistent across all waves.

Note that for this purpose codes (2) and (3) of JOBEND\$\$ for years 1985-1998 have been merged to one value, which is consistent to code (9) for years 1998 up to now. All three have been coded to (2) in COREJOBEND\$\$.
 The category "Other" contains all categories which were given only in some but not all years and the original different values for miscellaneous.

The respondents from the newly introduced migration sample were not asked about an occupational change, so no values were assigned to them.

Please also see JOBEND\$\$.

For more information, contact: SOEPmail (Tel. +49-30-89789-292)

ISEI\$\$ -- International Socio-Economic Index of Occupational Status

Var Label : ISEI\$\$ "International Socio-Economic Index of Occupational Status"
Value Label : ISEI\$\$ (-3)"No information for ISEI"
(-1)"No answer"
(-2)"Not applicable - not employed"
Var format : ISEI\$\$ (I2)
19\$\$ - Year : \$\$=84 - 13

Comment: This variable reflects the Standard International Socio-Economic Index of Occupational Status for all employed persons. The ISEI Index was developed in 1992 by Ganzeboom, De Graaf, Treiman, and De Leew based on information about income, education, and occupation. Technically, ISEI was created by scaling the ISCO88 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.

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, and 2008 - 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, ISEI\$\$ usually contains available information from the previous year for persons without a job change (persons with JOBCH\$\$-category (2) "employed, no change"). For some persons without a job change who updated the information on their current occupation without being asked, up-to-date information is available.

Please also see occupational prestige scores (SIOPSS\$, MPS\$\$) and occupational class (EGP\$\$).

Detailed description: Ganzeboom, H. B. G. / De Graaf, P .M. / Treiman, D. J. / De Leew, J.(1992): A Standard International Socio-Economic Index of Occupation Status, In: Social Science Research 21: 1-56

For more information, contact: Marco Giesselmann (Tel. +49-30-89789-503)

EGP\$\$ -- Erikson and Goldthorpe Class Category

Var Label : EGP\$\$ "Erikson and Goldthorpe Class Category"
Value Label : EGP\$\$ (-2)"not employed"
(-1)"no information on EGP level"
(1)"high service"
(2)"low service"
(3)"routine non-manual"
(4)"routine service-sales"
(5)"self-employed with employees"
(6)"self-employed without employees"
(8)"skilled manual"
(9)"semi-unskilled manual"
(10)"farm labor"
(11)"self-employed farmer"
(15)"not working - unemployed"
(18)"not working - pensioner"

Var format : EGP\$\$ (I2)
19\$\$ - Year : \$\$=84 - 13

Comment: This variable gives the occupational class for all employed persons. EGP\$\$ is derived from the Standard International Socio-Economic Index of Occupational Status (ISEI). Technically, the variable was created by scaling the ISCO-88 classification. In addition, it is based on information about income, education and occupation. The EGP Index was documented by Ganzeboom/Treiman in 1996 and revised in 2003.

The values for the variable range between 1 and 11; additional categories are (15) not working - registered unemployed and (18) not working - pensioner.

Non-working persons are only assigned to the category "not working - pensioner" if they are recipients of retirement pension or if they are recipients of widow's pension AND are older than 60 years. Moreover, if there is missing information on pension receipt, additional information from ARTKALEN (retrospective information from the activity calendar for the previous year) is used in the generation process to determine if a person was in retirement or early retirement (*Vorruhestand*) at the time of the interview. Hence, the category "not working - pensioner" in the most recent wave will be updated with retrospective information of the following wave. All other non-working persons are assigned to category (-2) "does not apply" as long as they are not registered as unemployed (category 15).

As information about supervisory status is only available from wave X (2007) on, it is not used to generate the corresponding EGP\$\$ category. Hence, the potential category (7) "Manual workers with supervisory status" is not assigned.

Annual information on the occupational position is used to generate the EGP-categories for the self-employed. In case no information on the number of employees is available, the EGP\$\$-categories (5) and (6) contain information on the firm size for self-employed persons.

Based on the new classification developed by Ganzeboom/Treiman (2003), several ISCO values were recoded in EGP\$\$ as follows:

- ISCO 2470 becomes EGP=1.

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

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, and 2008 - 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, EGP\$\$ usually contains available ISCO88-information from the previous year for persons without a job change. For some persons without a job change who updated the information on their current occupation without being asked, up-to-date information is available.

Please also see occupational status (ISEI\$\$) and occupational prestige scores (SIOP\$\$\$, MPS\$\$\$).

Detailed description: Ganzeboom, H. B. G. /Treiman, D. J. (1996): Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations In: Social Science Research 25: 201-239

Ganzeboom, H. B. G. /Treiman, D. J. (2003): Three Internationally Standardised Measures for Comparative Research on Occupational Status. In: Hoffmeyer-Zlotnik, J. H. P. Wolf, C. (eds.): Advances in Cross-National Comparison. A European Working Book for Demographic and Socio-Economic Variables. New York: Kluwer Academic/ Plenum Publishers. pp. 159-193.

For more information, contact: Marco Giesselmann (Tel. +49-30-89789-503)

SIOPSS\$ -- Treimans Standard Int. Occupation Prestige Score

Var Label : SIOPSS\$ **Treimans Standard Int. Occupation Prestige Score**
Value Label : SIOPSS\$ (-3)"no information for SIOPS"
(-2)"not employed"
(-1)"no answer"

Var format : SIOPSS\$ (I2)
19\$\$ - Year : \$\$=84..13

Comment: This variable gives the occupational prestige score index for all employed persons. SIOPSS\$ is based on ISCO-88 and was developed by Donald Treiman et al. The scale ranges from 6 to 78.

Not all employed persons are asked the 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, and 2008 - 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, SIOPSS\$ usually contains available information from the previous year for persons without a job change (persons with JOBCH\$\$-category (2) "employed, no change"). For some persons without a job change who updated the information on their current occupation without being asked, up-to-date information is available.

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

Detailed description: Ganzeboom, Harry B.G. and Donald Treiman (1996): Internationally comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations. In: Social Science Research, Vol. 25, 201-239

For more information, contact: Marco Giesselmann (Tel. +49-30-89789-503)

MPS\$\$ -- Magnitude-Prestige Scala - Wegener

Var Label : MPS\$\$ **"Magnitude-Prestige Scala - Wegener"**
Value Label : MPS\$\$ (-3)"no information for MPS"
(-2)"not working"
(-1)"no answer"

Var format : MPS\$\$ (I4)
19\$\$ - year : \$\$=84..13

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

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, and 2008 - 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, MPS\$\$ usually contains available information from the previous year for persons without a job change. For some persons without a job change who updated the information on their current occupation without being asked, up-to-date information is available.

Please also see occupational prestige scores (SIOP\$\$), occupational status (ISEI\$\$), and occupational class (EGP\$\$).

Detailed description: Wegener, Bernd (1988): Kritik des Prestiges, Opladen.

Frietsch, Rainer, and Heike Wirth (2001): Die Übertragung der Magnitude-Prestigeskala von Wegener auf die Klassifikation der Berufe. In: ZUMA Nachrichten 48 (Jg.25): 139-165.

For more information, contact: Marco Giesselmann (Tel. +49-30-89789-503)

PARTZ\$\$ -- Partner indicator

Var Label : PARTZ\$\$ **Partner indicator**
Value Label : PARTZ\$\$
(0)no partner
(1)spouse, registered partner
(2)partner
(3)probably spouse, registered partner
(4)probably partner

Var format : PARTZ\$\$ (I1)
19\$\$ - year : \$\$=84..13

COMMENT: Partner indicators have the purpose of defining singles, spouses (married and civil union), and partners (unmarried, no civil union) in SOEP households and thus enabling analyses on the couple level. The variable PARTZ\$\$ generated in this context reveals whether a person in a SOEP household has a partner in that household, and if so, the type of relationship existing between the partners. Relationships with persons outside the SOEP household are not covered by this variable.

To explain the codes:

Code 0 is assigned to all single persons living in households and those with partners outside the household. Codes 1 to 4 describe relationships. To assign Codes 1 and 2, the partnership has to be definable from the perspective of both partners unanimously. If conflicting information exists between partners, the codes 3 or 4 are assigned. If it is unclear whether an individual has no partner or whether she forms a couple with one other household member, we assign the code -1. Registered partnerships (civil unions) for same-sex couples were introduced in Germany in 2001. Though, registered partnerships are legally not equal to marriage, they are listed in the same category.

For more information, contact: Martin Kroh (Tel. +49-30-89789-678)

PARTNR\$\$ -- Person ID number of partner

Var Label : PARTNR\$\$ **Person ID number of partner**
Var format : PARTNR\$\$ (I6)
19\$\$ - year : \$\$=84..12

COMMENT: Partner indicators have the purpose of defining partner-relationships in SOEP households and thus to make possible analyses on the couple level.
If PARTZ\$\$ is coded 0, this person has no partner in the household. The variable PARTNR\$\$ is assigned the missing code of "-2" (=does not apply) for these persons.
If PARTZ\$\$ is coded 1, 2, 3 or 4, a partnership was defined and PARTNR\$\$ is then assigned the value of the unchanging person ID number (=PERSNR) of the partner. The assignment of the partner ID within households is based on four sources of information: A question in the person-file, that asks (unmarried) respondents to identify their partner in the household (bcpl3003 in 2013), the household matrix reported by the head of household at the beginning of the interview (bcstell in 2013), the partnership biography in the life-history 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, PARTNR\$\$ is self-consistent between two individuals.
For analyses of partner relationships, this information can be used to link all persons with their respective partners, and all information on both partners can also be stored in a common dataset.

For more information, contact: Martin Kroh (Tel. +49-30-89789-678)

\$FAMSTD -- Marital status in survey year

Var Label : \$FAMSTD **Marital status in survey year**
Value Label : \$FAMSTD
(-1)no answer
(-2)does not apply
(-3)implausible
(1)married
(2)married, separated
(3)single
(4)divorced
(5)widowed
(6)Spouse abroad
(7)Registered same-sex partnership
(8)Registered same-sex partnership, separated

Var format : \$FAMSTD (I1)
\$\$ - year : \$\$=84..12

COMMENT: Marital status is describing the institutional status of marriage at the time of the person interview. Marital status is based on information given by the respective person on his or her current relationship as well as on retrospective information about previous relationships asked in the biography questionnaire. Information on marital status when a child was born (provided in the biography information) is not used here, so contradicting information to BIOBIRTH might still be possible. For those whose partner was identified within the household, marital status is counter-checked with the information given by the partner. Where contradictions can be found, indication of the person information is compiled if reasonable. If no information is available, the indication by position related to head of household is deferred. Remaining contradictions are solved using information on marriage status when a child was born as well as future reports on a given relationship. Marital status is only available for people, who were interviewed. Data on marital status provided here is generated from and hence consistent with the corresponding BIOMARSY data, although value labels differ between \$FAMSTD and SPELLTYP. Note that marital status in a specific wave can be different between data distributions due to consistency checks using up-to-date information from following waves. Because of the newly introduced construction process of BIOMARSY, changes might be more than usual. Above, the partner indicator PARTZ\$\$ supplied in the \$PGEN data files as well might not match the information provided in \$FAMSTD in its entirety.

For more information, contact: Juliana Werneburg (Tel. +49-30-89789-217)

NATION\$\$ -- Citizenship - nationality

Var Label : NATION\$\$ **Citizenship - nationality**

Value Label : NATION\$\$ (1)Germany

(1) Germany	(55) Canada	(106) Montenegro
(2) Turkey	(56) New Zealand	(107) Belize
(3) Ex-Yugoslavia	(57) Tanzania	(108) Dominican Republic
(4) Greece	(58) Cyprus	(109) Nicaragua
(5) Italy	(59) Cuba	(110) Kenya
(6) Spain	(60) Iraq	(111) Libya
(7) Ex-GDR (Only Country Of Origin)	(61) Brazil	(112) Malta
(10) Austria	(62) Monaco	(113) Botswana
(11) France	(63) Hong Kong	(114) Haiti
(12) Benelux	(64) Peru	(115) Trinidad-Tobago
(13) Denmark	(65) Sri Lanka	(116) Luxembourg
(14) Great Britain	(66) Nepal	(117) Belgium
(15) Sweden	(67) Morocco	(118) Holland
(16) Norway	(68) China	(119) Croatia
(17) Finland	(69) Liechtenstein	(120) Bosnia-Herzegovina
(18) USA	(70) Iceland	(121) Macedonia
(19) Switzerland	(71) Ireland	(122) Slovenia
(20) Chile	(72) St. Lucia	(123) Slovakia
(21) Romania	(73) Moldavia	(124) Paraguay
(22) Poland	(74) Kazakhstan	(125) Guinea
(23) Korea	(75) Albania	(126) Kuwait
(24) Iran	(76) Lebanon	(127) Ivory Coast
(25) Indonesia	(77) Kyrgyzstan	(128) Malaysia
(26) Hungary	(78) Ukraine	(129) Samoa
(27) Bolivia	(79) Algeria	(130) Azerbaijan
(28) Portugal	(80) Mozambique	(131) Seychelles
(29) Bulgaria	(81) Egypt	(132) Belarus
(30) Syria	(81) Egypt	(133) Uruguay
(31) Czech Republic	(82) Tajikistan	(134) Bahamas
(32) Russia	(83) Vietnam	(135) Uganda
(33) Empty (was Kurdistan)	(84) Somalia	(136) Oman
(34) Mexico	(85) Pakistan	(137) Micronesia
(35) Argentina	(86) South Africa	(138) Mali
(36) Cap Verde Is.	(87) UAE	(139) Cameroon
(37) Benin	(88) El Salvador	(140) Kosovo-Albania
(38) Philippines	(89) Eritrea	(141) Georgia
(39) Israel	(90) Jordan	(142) Sudan
(40) Japan	(91) Turkmenistan	(143) Congo
(41) Australia	(92) Costa Rica	(144) Togo
(42) India	(93) Singapore	(145) Mongolia
(43) Afghanistan	(94) Burkina Faso	(146) Lithuania
(44) Thailand	(95) Zambia	(147) Chad
(45) Jamaica	(96) Ecuador	(148) Armenia
(46) Saudi Arabia	(97) Uzbekistan	(149) Kurdistan
(47) Ethiopia	(98) No Nationality	(150) Liberia
(48) Columbia	(99) Puerto Rico	(151) Yemen
(49) Ghana	(100) Laos	(152) Palaestina
(50) Bangladesh	(101) Estonia	(153) Freistaat Danzig
(51) Venezuela	(102) Angola	(154) Taiwan
(52) Tunisia	(103) Latvia	(155) Turkmenistan
(53) Mauritius	(104) Malaysia	
(54) Nigeria	(105) Namibia	

Var format : NATION\$\$ (I2)

19\$\$ - year : \$\$=84..13

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

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

\$PSBIL -- Diplomas/degrees from secondary/tertiary education

Var Label : \$PSBIL **Diplomas/degrees from secondary/tertiary education**
Value Label : \$PSBIL (1)Basic-track secondary school (9th grade)
(2)Intermediate-track secondary school (10th grade)
(3)Technical secondary school (12th grade)
(4)Academic-track secondary school (graduation from 13th grade)
(5)Other graduation diploma
(6)Left school without graduating
(7)Not yet graduated
Var format : \$PSBIL (I1)
\$ - year : 84..13

Comment: All respondents in all SOEP subsamples are asked about diplomas/degrees attained for completion of secondary/tertiary education (1984-1993 *blue* questionnaire; since 1994 *biographical questionnaire*) the first time they participate in SOEP. First: to generate this variable, the different diploma/degree categories provided for Subsamples B and D (see \$PSBILA) as well as C (see \$PSBIL0) 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.

Detailed description of the generation of Ypsbil, Ypsbila, Ypsbilo, Ypbbil01, Ypbbil02, Ypbbil03, Ypbbila Ypbbilo (Survey year 2008)

```
missing values all ()  
set case = uplow  
set len = none  
set blank = -2
```

```
+ get file = 'bild08.sav'.  
+ select if ynetto>=10 and ynetto<20.  
* desc var = all.  
+ sort cases by persnr  
+ save outfile = 'bil08.sav'  
/keep persnr hhnr yhhnr ynetto qpsbil to xbilzeit  
+ execute  
* finish  
+ get file = 'xpluecke.sav'  
+ compute plue = 1  
* desc var = all  
+ sort cases by persnr  
+ save outfile = 'bil_xplue.sav' /keep persnr plue xp1001 to xp1008  
+ execute  
* finish  
+ get file = 'yp.sav'  
* desc var = all  
+ sort cases by persnr  
+ save outfile = 'bil_yp.sav'  
/keep persnr yp16 to yp1706 yp8401 to yp8707  
* /keep persnr xp11 to xp1206 xp8301 to xp8607  
* /keep persnr wp05 to wp0605 wp7201 to wp7507 wpa201 to wpa409  
* /keep persnr vp08 to vp0905 vp8901 to vp9207  
* /keep persnr up07 to up0805 Up7001 to up7108  
* /keep persnr tp32 to tp3305 tp8501 to tp8608  
+ execute  
* finish  
+ get file = 'yjugend.sav'  
* desc var = all  
+ sort cases by persnr  
+ save outfile = 'bil_yjug.sav'  
/keep persnr yj26 to yj30 yj46 to yj48  
* /keep persnr xj26 to xj30 xj46 to xj48  
* /keep persnr wj26 to wj30 wj46 to wj48
```



```

*      /keep persnr vj25 to vj27 vj44 to vj45e02
*      /keep persnr uj25 to uj27 uj44 to uj45e02
*      /keep persnr tj25 to tj27 tj44 to tj45e02
+ execute
* finish
+ get   file = 'ylela.sav'
*      desc var = all
+ sort cases by persnr
+ save outfile = 'bil_ylela.sav'
      /keep persnr yb3701 to yb5408
*      /keep persnr xb3701 to xb5408
*      /keep persnr wb3701 to wb5408
*      /keep persnr vb3701 to vb5408
*      /keep persnr ub3701 to ub5408
*      /keep persnr tb3701 to tb5408
+ execute
* finish

match files  file = 'bil08.sav'
             /file = 'bil_yp.sav'
             /file = 'bil_yjug.sav'
             /file = 'bil_ylela.sav'
             /file = 'bil_xplue.sav'
             /by persnr

+ fre  Xpsbil Xpsbila Xpsbilo
      Xpbbil01 to Xpbbil03 Xpbbila Xpbbilo

+ select if Ynetto>=10 and Ynetto<20 or plue = 1
* select if Wnetold=1 or Wnetold=5 or Wnetold=7 or plue = 1

* desc var = all
* finish
missing values all ( )
+ desc var = all
* finish
* -----
* ~~~~~ 1. Vorbesetzen der Variablen ~~~~~
* -----
do repeat
  a = Ypsedu Ypbedu01 to Ypbedu03
      psedu pbedu01 to pbedu03
+   compute a = -1
end repeat
do repeat
  b = Ypseduo Ypbeduo Ypsedua Ypbedua
      pseduo pbeduo psedua pbedua
+   compute b = -2
end repeat

+ do repeat
  a = xpsbil  wpsbil  vpsbil  upsbil  tpsbil  spsbil  rpsbil
      qpsbil
  /b = xpbbil01 wpbbil01 vpbbil01 upbbil01 tpbbil01 spbbil01 rpbbil01
      qpbbil01
  /c = xpbbil02 wpbbil02 vpbbil02 upbbil02 tpbbil02 spbbil02 rpbbil02
      qpbbil02
  /d = xpbbil03 wpbbil03 vpbbil03 upbbil03 tpbbil03 spbbil03 rpbbil03
      qpbbil03
  /e = xpsbila wpsbila vpsbila upsbila tpsbila spsbila rpsbila
      qpsbila
  /f = xpbbila wpbbila vpbbila upbbila tpbbila spbbila rpbbila
      qpbbila
  /g = xpsbilo wpsbilo vpsbilo upsbilo tpsbilo spsbilo rpsbilo
      qpsbilo
  /h = xpbbilo wpbbilo vpbbilo upbbilo tpbbilo spbbilo rpbbilo
      qpbbilo
+   if (psedu < 0 and a gt 0 and a < 7) psedu = a
+   if (pbedu01 < 0 and b gt 0) pbedu01 = b
+   if (pbedu02 < 0 and c gt 0) pbedu02 = c
+   if (pbedu03 < 0 and d gt 0) pbedu03 = d
+   if (psedua < 0 and e gt 0) psedua = e
+   if (pbedua < 0 and f gt 0) pbedua = f
+   if (pseduo < 0 and g gt 0) pseduo = g
+   if (pbeduo < 0 and h gt 0) pbeduo = h
+   end repeat

* *****
* ----- Update _psedu -----
* *****

```

```

* -----
* ~~~~ 2. in schulischer Ausbildung ~~~~~
* -----
+         if (Yp1701 ge 1)                Ypsedu = 7
+         if (Yb3702 eq 1)                Ypsedu = 7
+         if (Yj26 ge 1 and Yj26 le 4)    Ypsedu = 7
*
*         if (_pa201 eq 1)                Ypsedu = 7
* -----
* ~~~~ 3. Ueberschreiben mit Vorjahresinformationen ~~~~~
* -----
.         if (psedu > 0 and psedu < 7)    Ypsedu = psedu
.
.         if (Xp1001 ge 1 and Xp1001 le 4) Ypsedu = Xp1001
. do if (Xp1001 eq 5)
.   if (Ypsedu lt 1 or Ypsedu gt 4)      Ypsedu = 5
. end if
* -----
* ~~~~ 4. Ueberschreiben mit aktuellen Angaben ~~~~~
* -----
* [if (_xxx_ eq 1)                Ypsedu = 7]
+   if (Yb42 eq 1)                Ypsedu = 6
+   if (Yb42 eq 6)                Ypsedu = 5
+   if (Yb42 eq 2)                Ypsedu = 1
+   if (Yb42 eq 3)                Ypsedu = 2
+   if (Yb42 eq 4)                Ypsedu = 3
+   if (Yb42 eq 5)                Ypsedu = 4
+
+   if (Yj28 eq 1)                Ypsedu = 6
+   if (Yj28 eq 5)                Ypsedu = 5
+   if (Yj28 eq 2)                Ypsedu = 1
+   if (Yj28 eq 3)                Ypsedu = 2
+   if (Yj28 eq 4)                Ypsedu = 3
*
*   if (_pa202 ge 1)              Ypsedu = _pa202
* -----
* ~~~~ 4.1 Generieren von _psedua / _pseduo ~~~~~
* -----
*
*         if (psedua > 0)                Ypsedua = psedua
+         if (value(psedua) ge -2)       Ypsedua = psedua
.         if (Yb40 ge 1)                 Ypsedua = Yb40
.         if (Ypsedua < 0 and (value(Yb40) = -1)) Ypsedua = -1
*
*         if (Ypa205 ge 1)               Ypsedua = Ypa205
*         if (Ypsedua < 0 and (value(Ypa205) = -1)) Ypsedua = -1
+
+         if (pseduo > 0)                Ypseduo = pseduo
+ do if (Yb38 = 2)
+   if (Yb42 eq 1)                Ypseduo = 5
+   if (Yb42 eq 6)                Ypseduo = 4
+   if (Yb42 eq 2)                Ypseduo = 1
+   if (Yb42 eq 3)                Ypseduo = 2
+   if (Yb42 eq 4)                Ypseduo = 3
+   if (Yb42 eq 5)                Ypseduo = 3
+ end if
*
*         if (_pa203 eq 5)              Ypseduo = 5
*         if (_pa203 eq 4)              Ypseduo = 4
*         if (_pa203 eq 1)              Ypseduo = 1
*         if (_pa203 eq 2)              Ypseduo = 2
*         if (_pa203 eq 3)              Ypseduo = 3
* -----
* ~~~ 5. Ueberschreiben mit neu erworbenen Abschlussen ~~~
* -----
.         if (Yp8701 ge 1 and Yp8701 le 4) Ypsedu = Yp8701
. do if (Yp8701 eq 5)
.   if (Ypsedu lt 1 or Ypsedu gt 4)      Ypsedu = 5
. end if
* -----
* ~~~ 6. Integration von anderen Abschlussen (_psedua/o) ~
* -----

```

```

.         if ((Ypsedu < 0) and (Ypsedua = 1)) Ypsedu = 6
.         if ((Ypsedu < 0) and (Ypsedua > 1)) Ypsedu = 5
.         if ((Ypsedu = 6) and (Ypsedua > 1)) Ypsedu = 5

+         if      (Ypseduo > 0 and Ypsedu < 0) Ypsedu = Ypseduo
+ do if      (Ypseduo > 0 and Ypsedu > 4)
+         if      (Ypseduo = 1 and Ypsedu > 4) Ypsedu = 1
+         if      (Ypseduo = 2 and Ypsedu > 4) Ypsedu = 2
+         if      (Ypseduo = 3 and Ypsedu > 4) Ypsedu = 4
+         if      (Ypseduo = 4 and Ypsedu > 4) Ypsedu = 5
+     end if

* *****
* ----- Update _pbedu01 to _pbedu03 -----
* *****

* -----
* ~~~~ a. Ueberschreiben mit Xorjahresinformationen ~~~~~
* -----

+     if (pbedu01 > 0 )           Ypbedu01 = pbedu01
+     if (pbedu02 > 0 )           Ypbedu02 = pbedu02
+     if (pbedu03 > 0 )           Ypbedu03 = 1

.     if      (Xp1006 eq 1)           Ypbedu01 = 1
.     if      (Xp1006 eq 2 or Xp1006 eq 3) Ypbedu01 = 2
.     if      (Xp1006 eq 4)           Ypbedu01 = 4
.     if      (Xp1006 eq 5)           Ypbedu01 = 5
. do if      (Xp1006 eq 6 or Xp1006 eq 7)
.     if      (Ypbedu01 lt 1 )       Ypbedu01 = 6
. end if
. do if      (Xp1003 ge 1)
.     if      ((Xp1003 gt Ypbedu02) or Ypbedu02 lt 1) Ypbedu02 = Xp1003
. end if

* -----
* ~~~~ b. Ueberschreiben mit aktuellen Angaben ~~~~~
* -----

.     if (Yb4713 = 1)           Ypbedu01 = 6
.     if (Yb4701 = 1)           Ypbedu01 = 1
.     if (Yb4703 = 1)           Ypbedu01 = 2
.     if (Yb4705 = 1)           Ypbedu01 = 4
.     if (Yb4707 = 1)           Ypbedu01 = 5

.     if (Yb4709 = 1)           Ypbedu02 = 1
.     if (Yb4711 = 1)           Ypbedu02 = 2
+ do if (Yb38 = 2 and (Yb48 ge 1950 and Yb48 le 1991))
+     if (Yb4709 = 1)           Ypbedu02 = 4
+     if (Yb4711 = 1)           Ypbedu02 = 5
+ end if

.     if (Yb46 = 2)           Ypbedu03 = 1

.     if (Yj4706 = 1)           Ypbedu01 = 1
.     if (Yj4708 = 1)           Ypbedu01 = 2

.     if (Yj46 = 2)           Ypbedu03 = 1

*     if (_pa409 = 1)           Ypbedu01 = 6
*     if (_pa401 = 1 or Ypa402 = 1) Ypbedu01 = 1
*     if (_pa403 = 1 or Ypa404 = 1) Ypbedu01 = 2
*     if (_pa405 = 1)           Ypbedu01 = 4
*     if (_pa406 = 1)           Ypbedu01 = 5

*     if (_pa407 = 1)           Ypbedu02 = 1
*     if (_pa408 = 1)           Ypbedu02 = 2

*     if (_pa3 = 2)           Ypbedu03 = 1

* -----
* ~~~~ b.1 Generieren von _pbedua / _pbeduo ~~~~~
* -----

+         if (pbedua > 0)           Ypbedua = pbedua

.         if (Yb5005 eq 1)           Ypbedua = 5
.         if (Yb5001 eq 1)           Ypbedua = 1

```

```

.      if (Yb5002 eq 1)      Ypbedua = 2
.      if (Yb5003 eq 1)      Ypbedua = 3
.      if (Yb5004 eq 1)      Ypbedua = 4

*      if (psample eq 7)      Ypbedua = -3

+      if (pbeduo > 0)        Ypbeduo = pbeduo

+ do if (Yb38 = 2 and (Yb48 ge 1950 and Yb48 le 1991))
.      if (Yb4713 = 1)        Ypbeduo = 4
.      if (Yb4701 = 1)        Ypbeduo = 1
.      if (Yb4703 = 1)        Ypbeduo = 1
.      if (Yb4705 = 1)        Ypbeduo = 2
.      if (Yb4707 = 1)        Ypbeduo = 3
+ end if

* -----
*   ~~~ c. Ueberschreiben mit neu erworbenen Abschlussen ~~~
* -----

.      if (Yp8706 eq 1)        Ypbedu01 = 1
.      if (Yp8706 eq 2 or Yp8706 eq 3) Ypbedu01 = 2
.      if (Yp8706 eq 4)        Ypbedu01 = 4
.      if (Yp8706 eq 5)        Ypbedu01 = 5
.      do if (Yp8706 eq 6 or Yp8706 eq 7)
.      if (Ypbedu01 lt 1)      Ypbedu01 = 6
.      end if
.      do if (Yp8702 ge 1)
.      if ((Yp8702 gt Ypbedu02) or Ypbedu02 lt 1) Ypbedu02 = Yp8702
.      end if

* -----
*   ~~~~ d. in beruflicher Ausbildung ~~~~~
* -----

+ if (Yp1705 ge 1)            Ypbedu03 = 2

+ if (Yj4701 = 1)            Ypbedu03 = 2
+ if (Yj4703 = 1)            Ypbedu03 = 2
+ if (Yj4705 = 1)            Ypbedu03 = 2
+ if (Yj4707 = 1)            Ypbedu03 = 2
+ if (Yj4709 = 1)            Ypbedu03 = 2

+ if (Yp1702 ge 1)            Ypbedu03 = 3

* -----
*   ~~~ e. Integration und Missings ~~~~~
* -----

.      do if ( Ypbedu01 < 0)
.      if (Yb5002 =1 or Yb5003 =1 or Yb5005 =1) Ypbedu01 = 6
.      end if
.      do if (Ypbedu02 < 0)
.      if (Ypbeduo = 3)        Ypbedu02 = 4
.      end if
.      do if (Ypbedu02 < 0)
.      if (Yb5004 = 1)        Ypbedu02 = 3
.      end if
.      do if (Ypbedu03 < 0)
.      if (Yb49 = 2)          Ypbedu03 = 1
+      if (Yb46 = 2)          Ypbedu03 = 1
.      end if

* -----
.      if (((Ypbedu03 = -1) or (Ypbedu03 = 1)) and
.      ((Ypbedu01 ge 1) or (Ypbedu02 ge 1))) Ypbedu03 = -2
.      if (((Ypbedu01 = -1) and ((Ypbedu02 ge 1) or (Ypbedu03 ge 1)))
.      Ypbedu01 = -2
.      if (((Ypbedu02 = -1) and ((Ypbedu01 ge 1) or (Ypbedu03 ge 1)))
.      Ypbedu02 = -2

* -----
* -----
* -----
* -----

+ compute      Ypsbil      = Ypsedu
+ compute      Ypbbil01    = Ypbedu01
+ compute      Ypbbil02    = Ypbedu02
+ compute      Ypbbil03    = Ypbedu03
+ compute      Ypsbila     = Ypsedua

```

+ compute Ypbbila = Ypbedua
+ compute Ypsbilo = Ypseduo
+ compute Ypbbilo = Ypbeduo
* -----

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

\$PSBILO -- Secondary school degree/diploma East Germany

Var Label : \$PSBILO **Secondary school degree/diploma East Germany**
Value Label : \$PSBILO (1)completion of 8th grade
(2)completion of 10th grade
(3)college entrance exam
(4)other degree/diploma
(5)dropout, no degree/diploma
Var format : \$PSBILO (I1)
\$ - year : 84..13

Comment: As a supplement to the variable \$PSBIL the highest secondary school degree/diploma in East Germany is provided as a separate variable and updated if necessary for 1991. Since 1992, secondary degrees/diplomas are asked only in the West German version. New SOEP respondents are also asked about secondary degrees/diplomas obtained in the former GDR; and for old respondents, the same codes are carried forward.

Detailed description (cf. \$psbil)
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

\$PSBILA -- Secondary school degrees/diplomas abroad

Var Label : \$PSBILA **Secondary school degrees/diplomas abroad**
Value Label : \$PSBILA (1)secondary school, no degree/diploma attained
(2)secondary school, degree/diploma attained
(3)vocational school
Var format : \$PSBILA (I1)
\$ - year : 84..13

Comment: As a supplement to the \$PSBIL, this variable provides annually updated data on the highest secondary school degree/diploma attained abroad.

Detailed description (cf. \$psbil)
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

\$PBBIL01 -- Vocational degree attained

Var Label : \$PBBIL01 **Vocational degree attained**
Value Label : \$PBBIL01 (1)apprenticeship
(2)vocational school
(3)health care school (since 2001 integrated into (4))
(4)technical school
(5)civil service training
(6)other training
Var format : \$PBBIL01 (I1)
\$ - year : 84..13

Comment: All respondents in all subsamples are asked about vocational degrees attained the first time they participate in SOEP (1984-1993 *blue* questionnaire; since 1994 *biographical* questionnaire). To generate the variable, the different vocational degrees for Subsamples B and D (cf. \$PBBILA) as well as C (cf. \$PBBILO) are integrated into the West German vocational degree categories (Subsample A). The categories that originally each constituted individual variables are combined to make them compatible with the annual question about changes in vocational degrees attained, and this data is updated annually.

Detailed description (cf. \$psbil)
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

\$PBBILO -- Vocational degree attained - East

Var Label : \$PBBILO **Vocational degree attained - East**
Value Label : \$PBBILO (1)vocational training
(2)master craftsman
(3)engineering, technical degree
(4)other training
Var format : \$PBBILO (I1)
\$ - year : 84..13

Comment: To supplement the variable \$PBBIL01 the highest secondary school degree/diploma in East Germany is provided as a separate variable and updated if necessary for 1991. Since 1992 only the West German version has been used for new vocational degrees. For new SOEP respondents, vocational degrees attained in the former GDR are asked as well; for old respondents, the same codes are carried forward. From 2002 on, the questionnaire was expanded and revised, but this led to an operationalization involving more assumptions on the vocational degrees attained in the GDR; (from 2002 on, Code 3 is also listed as the additional category Code 4 in the integrated variables \$PBBIL03 if this degree has not been replaced by a more recently attained, higher-level university or college degree).

Detailed description (cf. \$psbil)
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

\$PBBIL02 -- Completed college education

Var Label : \$PBBIL02 **Completed college education**
Value Label : \$PBBIL02 (1)technical college
(2)university, technical university
(3)college abroad
(4)engineering, technical school (East)
(5)university (East)
Var format : \$PBBIL02 (I1)
\$ - year : 84..13

Comment: 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.

Detailed description (cf. \$psbil)
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

\$PBBIL03 -- No vocational degree

Var Label : \$PBBIL03 **No vocational degree**
Value Label : \$PBBIL03 (1)no vocational degree
(2)apprenticeship
(3)university
Var format : \$PBBIL03 (I1)
\$ - year : 84..13

Comment: In connection with the question about vocational degrees (\$PBBIL01 and \$PBBIL02), all first-time respondents to all subsamples are explicitly asked whether they (still) *do not* possess a vocational degree. In the subsequent years, this data is carried forward or updated. The variable has the Missing Value Code -2 (does not apply) if one of the other two variables on vocational degree has a positive value.

Detailed description (cf. \$psbil)
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

\$PBBILA -- Vocational degree abroad

Var Label : \$PBBILA **Vocational degree abroad**
Value Label : \$PBBILA (1)on-the-job training
(2)vocational training
(3)vocational school
(4)college
(5)other
Var format : \$PBBILA (I1)
\$ - year : 84..13

Comment: As a supplement to the variable \$PBBIL01, this variable gives
(and updates) the highest-level vocational degree attained
abroad.

Detailed description (cf. \$psbil)
For more information, contact: Peter Krause (Tel. +49-30-89789-690)

\$BILZEIT -- Amount of education or training (in years)

Var Label : \$BILZEIT **Amount of education or training (in years)**
Var format : \$BILZEIT (I4)
\$ - year : 84..13

Comment: The following statements describe the standard computation for schooling (including years of secondary vocational education).
As can be seen, the code is not very differentiated. For example, special schools for health care professions and other kinds of specialized schools are all included in the "technical school" label. However, in Germany, this code is the most commonly used one when earnings functions based on human capital theory are estimated.

\$BILZEIT is now computed for all samples.

Computation The \$BILZEIT variables are computed using the education variables provided by the \$PGEN-files. The computation code is as follows:

```
school = -1 *** preset of intermediate schooling variable****
occupa = -1 *** preset of intermediate occupational variable***

*** computation for east-germany (mainly samples C and D) ***

ifthen (xPSBILO ge 1 and school eq -1)
+ if (xPSBILO eq 1) school = 1
+ if (xPSBILO eq 2) school = 2
+ if (xPSBILO eq 3) school = 4
+ if (xPSBILO eq 4) school = 2
+ if (xPSBILO eq 5) school = 0

ifthen (xPBBILO ge 1 and occupa eq -1)
+ if (xPBBILO eq 1) occupa = 1
+ if (xPBBILO eq 2) occupa = 2
+ if (xPBBILO eq 3) occupa = 3
+ if (xPBBILO eq 4) occupa = 2

*** computation for foreigners (sample B and D) **
*** finished in their country of origin **

ifthen (xPSBILA ge 1 and school eq -1)
  if (xPSBILA eq 1) school = 0
  if (xPSBILA eq 2) school = 1
  if (xPSBILA eq 3) school = 2
end if
ifthen (xPBBILA ge 1 and occupa eq -1)
  . if (xPBBILA eq 1) occupa = 0
  . im (xPBBILA ge 2) occupa = 1
end if

*** All samples with finished education in Germany: ****

if (xPSBIL eq 6) school = 0
if (xPSBIL ge 1 and xPSBIL le 4) school = xPSBIL
if (xPSBIL eq 5) school = 2

if (xpsbil eq 5 and xpsbila eq 2) school = 1
if (xpsbil eq 5 and xpsbila eq 3) school = 2

if (xPBBIL03 ge 1) occupa = 0
if (xPBBIL01 eq 1 or xPBBIL01 eq 5) occupa = 1
if ((xPBBIL01 ge 2 and xPBBIL01 le 4) or xPBBIL01 eq 6) occupa = 2
if (xPBBIL02 eq 1) occupa = 3
if (xPBBIL02 ge 2) occupa = 4
```

```

c ** years of education 0          *****
c ** years of education=          *****
c ** years of schooling + years of occupational training ***
c ** schooling                    *
c **   no degree                  = 7   years   *
c **   lower school degree        = 9   years   *
c **   intermediary school        = 10  years   *
c **   degree for a professional coll. = 12 years   *
c **   high school degree         = 13  years   *
c **   other                      = 10  years   *
c ** additional occupational training (includes universities) *
c **   apprenticeship             = 1.5 years   *
c **   technical schools (incl. health) = 2   years   *
c **   civil servants apprenticeship = 1.5 years   *
c **   higher technical college    = 3   years   *
c **   university degree          = 5   years   *
c ** *****

```

```

xBILZEIT = -1          *** years of education ***
if (school eq 0)      xBILZEIT = 7
if (school eq 1)      xBILZEIT = 9
if (school eq 2)      xBILZEIT = 10
if (school eq 3)      xBILZEIT = 12
if (school eq 4)      xBILZEIT = 13
if (xpsbil eq 7)      xbilzeit = -2 /**noch in Schule
ifthen (xBILZEIT ge 7)
if (occupa eq 1)      xBILZEIT = xBILZEIT + 1.5
if (occupa eq 2)      xBILZEIT = xBILZEIT + 2
if (occupa eq 3)      xBILZEIT = xBILZEIT + 3
if (occupa eq 4)      xBILZEIT = xBILZEIT + 5
endif

```

c *** end of program *****

Detailed description Helberger, Christof (1988): Eine Überprüfung der Linearitätsannahme der Humankapitaltheorie. In: H.-J. Bodenhöfer (ed.): Bildung, Beruf, Arbeitsmarkt, pp. 151-170, Berlin.
 Schwarze, Johannes (1991): Ausbildung und Einkommen von Männern - Einkommensfunktionsschätzungen für die ehemalige DDR und die Bundesrepublik Deutschland. In: Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, (24), pp. 63-69.

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

\$ISCED -- Highest degree/diploma attained, ISCED-1997 classification

Var Label : \$ISCED Highest degree/diploma attained, ISCED-1997 classification

Value Label : \$ISCED
 (0)'in school'
 (1)'inadequately'
 (2)'general elementary'
 (3)'middle vocational'
 (4)'vocational + Abi'
 (5)'higher vocational'
 (6)'higher education'

Var format : \$ISCED (I1)
 \$ - year : 84..13

Comment: To make the educational degrees and diplomas attained in different countries comparable, for all respondents an educational variable (\$ISCED) is generated retroactively from 1984 on using the international classification scheme ISCED-1997 (International Standard Classification of Education). It creates the highest degree/diploma attained, taking into account degrees and diplomas attained in both general schooling and in vocational and university education. Here the higher-level vocational and university override lower-level school diplomas. Persons who, for example, have no values for the variables on secondary school degrees/diplomas but state that they have a university degree are placed in the highest ISCED category. Please note that, due to a lack of more detailed information on tertiary degrees -- in particular on promotion -- we include all tertiary degrees in our ISCED category 6. Thus, the ISCED variable provided here is not comparable one-to-one with the ISCED levels as defined by the OECD, since we have included the original ISCED level 5A in our ISCED category 6. See below for more details.

Computation: The \$ISCED variables are computed using the education variables provided by the \$PGEN-files. For this we use the variables on secondary degrees/diplomas (\$PSBIL) and secondary degrees/diplomas abroad (\$PSBILA), and the occupational education variables "vocational degree"(\$PBBIL01), "university degree" (\$PBBIL02) and "vocational degree abroad" (\$PBBILA). We refrained from integrating the GDR-specific educational degrees/diplomas (\$PSBILO und \$PBBILO) since only the West German version of the question has been asked since 1992.

The computation algorithm can be summarized as follows:

```

isced=-1 if netold==1 | netold==5 | netold==7          *population
isced=0  if psbil==7                                     *0:in school
isced=1  if psbil==5 | psbil==6 | psbila==1           *1:inadequately
isced=2  if psbil==1 | psbil==2 | psbila==2           *2:gen.element.
isced=3  if psbil==3 | psbil==4 | psbila==3 |         *3:middle voc.
          pbbil01==1 | pbbil01==2 | pbbil01==6 |
          pbbila==2 | pbbila==3
isced=4  if (psbil==3 | psbil==4 | psbila==3) &      *4:voc. + Abi
          (pbbil01==1 | pbbil01==2 | pbbil01==6 |
          pbbila==2 | pbbila==3)
isced=5  if pbbil01==3 | pbbil01==4 |                 *5:higher voc.
          (pbbil01==5 & (psbil==3 | psbila==3 ))
isced=6  if pbbil02==1 | pbbil02==2 | pbbil02==3 |   *6:higher educ.
          pbbil02==4 | pbbil02==5 | pbbila==4
  
```

ISCED SOEP	PSBIL Schulabschluss	PSBILA Schule im Ausland		PBBIL01 Berufliche Ausbildung	PBBILA Berufsbildung im Ausland	PBBIL02 Hochschul- abschluss	ISCED97 – OECD
0	7.Noch kein Abschluss						0 noch kein Abschluss
1	5.Anderer Abschluss 6.Ohne Abschluss verlassen	1.Pflichtschule ohne Abschluss					1 ohne Abschluss verlassen
2	1.Hauptschul- abschluss 2.Realschulabschluss	2.Pflichtschule mit Abschluss					2 Haupt/Real- schulabschluss
3	3.Fachhochschulreife 4.Abitur	3.Weiterfuehrende Schule		1.Lehre 2.Berufsfachschule 6.Sonstiger Abschluss	2.Betriebliche Ausbildung 3.Berufsbildende Schule		Beruflicher/Real- schulabschluss oder (Fach)Abitur
4	3.Fachhochschulreife 4.Abitur	3.Weiterfuehrende Schule	&	1.Lehre 2.Berufsfachschule 6.Sonstiger Abschluss	2.Betriebliche Ausbildung 3.Berufsbildende Schule		Beruflicher Abschluss nach 4 absolvierter allgemeinb. Schule
5				3.Schule Gesundheitsw.(-99) 4.Fachschule, Meister			höherer 5Bberuflicher Abschluss
	3.Fachhochschulreife 4.Abitur	3.Weiterfuehrende Schule	&	5.Beamtenausbildung			
6					4.Hochschule	1.Fachhochschule 2.Universitaet, TH 3.Hochschule im Ausland 4.Ingenieur/Fach- schule (Ost) 5.Hochschule (Ost)	5A FH oder Universität + 6 Promotion

Detailed description OECD (1999): Classifying Educational Programmes: Manual for ISCED-97
Implementation in OECD Countries. Paris 1999.

For more information, contact: Pia Schober (Tel. +49-30-89789-691)

\$CASMIN -- Highest degree/diploma according to CASMIN classification

Var Label : \$CASMIN Highest degree/diploma according to CASMIN classification

Value Label : \$CASMIN
 (0)'(0) in school'
 (1)'(1a) inadequately completed '
 (2)'(1b) general elementary school'
 (3)'(1c) basic vocational qualification'
 (4)'(2a) intermediate general qualification'
 (5)'(2b) intermediate vocational'
 (6)'(2c_gen) general maturity certificate'
 (7)'(2c_voc) vocational maturity certificate'
 (8)'(3a) lower tertiary education'
 (9)'(3b) higher tertiary education'

Var format : \$CASMIN (I1)
 \$ - year : 84..13

Comment: As an alternative to \$ISCED, a second educational variable is generated (\$CASMIN) that also enables comparison with international educational degrees/diplomas. Based on the modified CASMIN classification scheme (Comparative Analysis of Social Mobility in Industrial Nations), this variable has been computed retroactively from 1984 on for all respondents. Taken into account are both secondary-level and university/college-level degrees and diplomas. As with \$ISCED, the higher-level occupational degrees override the lower-level secondary school degrees.

Computation: The \$CASMIN variables are computed using the education variables provided by the \$PGEN-files. For this we use the variables on secondary degrees/diplomas (\$PSBIL) and secondary degrees/diplomas abroad (\$PSBILA), and the occupational education variables "vocational degree"(\$PBBIL01), "university degree" (\$PBBIL02) and "vocational degree abroad" (\$PBBILA). We refrained from integrating the GDR-specific educational degrees/diplomas (\$PSBILO und \$PBBILO) since only the West German version of the question has been asked since 1992.

The computation algorithm can be summarized as follows:

```
casmin=-1 if netold==1 | netold==5 | netold==7 *population
casmin= 0 if psbil==7 *0:0 in school
casmin= 1 if psbil==5 | psbil==6 | psbila==1 *1:1a inadequately completed
casmin= 2 if psbil==1 | psbila==2 *2:1b general elementary school
casmin= 3 if (psbil==1 | psbil==5 | psbil==6 | psbila==2) &
      (pbbila==2 | pbbila==3 |
      (pbbil01>=1 & pbbil01<.)) *3:1c basic vocational
casmin= 4 if (psbil==2 | psbila==3) *4:2b intermediate general
casmin= 5 if (psbil==2 | psbila==3) &
      (pbbila==2 | pbbila==3 |
      (pbbil01>=1 & pbbil01<.)) *5:2a intermediate vocational
casmin= 6 if (psbil==3 | psbil==4) *6:2c_gen general maturity cer.
casmin= 7 if (psbil==3 | psbil==4) &
      (pbbila==2 | pbbila==3 |
      (pbbil01>=1 & pbbil01<.)) *7:2c_voc vocational maturity
casmin= 8 if pbbil02==1 *8:3a lower tertiary education
casmin= 9 if (pbbil02>=2 & pbbil02<=5) | pbbila==4 *9:3b higher tertiary education
```

CASMIN	PSBIL Schulabschluss	PSBILA Schule im Ausland		PBBIL01 Berufliche Ausbildung	PBBILA Berufsbildung im Ausland	PBBIL02 Hochschulabschluss
0 – 0	7. Noch kein Abschluss					
1 – 1a	5. Anderer Abschluss 6. Ohne Abschluss verlassen	1. Pflichtschule ohne Abschluss				
2 – 1b	1. Hauptschulabschluss	2. Pflichtschule mit Abschluss				
3 – 1c	1. Hauptschulabschluss 5. Anderer Abschluss 6. Ohne Abschluss verlassen	2. Pflichtschule mit Abschluss	&	1-6. Beruflicher Bildungsabschluss	2. Betriebl. Ausbildung 3. Berufsbild. Schule	
4 – 2a	2. Realschulabschluss	3. Weiterfueh- rende Schule				
5 – 2b	2. Realschulabschluss	3. Weiterfueh- rende Schule	&	1-6. Beruflicher Bildungsabschluss	2. Betriebl. Ausbildung 3. Berufsbild. Schule	
6 – 2c_gen	3. Fachhochschulreife 4. Abitur					
7 – 2c_voc	3. Fachhochschulreife 4. Abitur		&	1-6. Beruflicher Bildungsabschluss	2. Betriebl. Ausbildung 3. Berufsbild. Schule	
8 – 3a						1. Fachhochschule
9 – 3b					4. Hochschule	2. Universitaet, TH 3. Hochschule im Ausland 4. Ingenieur/Fachs. Ost 5. Hochschule (Ost)

Detailed description The original version is described in: König, W./Lüttinger, P./Müller, W. (1988): A Comparative Analysis of the Development and Structure of Educational Systems. Methodological Foundations and the Construction of a Comparative Educational Scale. CASMIN Working Paper No. 12. Mannheim: Universität Mannheim.
For the modified version see: Brauns, H./Steinmann, (1999): Educational Reform in France, West-Germany and the United Kingdom: Updating the CASMIN Educational Classification. In: ZUMA Nachrichten, Jg. 23, H. 44, pp. 7-44.

For more information, contact: Pia Schober (Tel. +49-30-89789-691)

MONTH\$\$ -- Month of interview

Var Label : MONTH\$\$ **Month of interview**

Value Label : MONTH\$\$ (1)'January'
(2)'February'
(3)'March'
(4)'April'
(5)'May'
(6)'June'
(7)'July'
(8)'August'
(9)'September'
(10)'October'
(11)'November'
(12)'December'

Var format : MONTH\$\$ (I1)

19\$\$ - year : \$\$=84..13

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

For more information, contact: Jürgen Schupp (Tel. +49-30-89789-238)

MODE\$\$ -- Interview method

```
Var Label      :  MODE$$      Interview method
Value Label    :  MODE$$      (100)with interviewer assistance
                                     (110)oral interview
                                     (120)written questionnaire (without interviewer)'
                                     (130)mixed form (with/without interviewer
assistance)
                                     (131)written questionnaire (with interviewer
assistance)'
                                     (132)oral and written
                                     (133)proxy
                                     (134)third person present
                                     (135)without third person
                                     (140)CAPI - Wave 0 onwards
                                     (200)with telephone assistance
                                     (210)written, by mail
                                     (220)telephone interview
Var format     :  MODE$$      (I2)
19$$ - year   :  $$=84..13
```

Comment: The interview method is generated via the answers to the questions in the individual questionnaire. Missing answers are filled in from the \$pbrutto files.

For more information, contact: Jürgen Schupp (Tel. +49-30-89789-238)

LABGRO\$\$ -- Current gross labor income in euros (generated)

Var Label : LABGRO\$\$ **Current gross labor income in euros (generated)**

Var format : LABGRO\$\$ (F2)

19\$\$ - year : \$\$=84..13

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

The original variables coming from the \$P-files and are:
ap3301,bp4301,cp5201,dp4401,ep4401,fp4501,gp4301,hp5401
ip5401,jp5401,kp6401,lp5301,mp4701,np5401,op4501,pp6001
qp5601,rp5701,sp5801,tp7601,up5901,vp7101,wp5901,xp7301
yp6801,zp7201,bap6101

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

IMPGRO\$\$ -- Imputation flag for LABGRO\$\$

Var Label : IMPGRO\$\$ **Imputation flag for LABGRO\$\$**

Var format : IMPGRO\$\$ (I1)

19\$\$ - year : \$\$=84..13

Comment:

The variable IMPGRO\$\$ designates imputations of item-nonresponse in the variable LABGRO\$\$ (current gross labor income). IMPGRO\$\$ can take the value
0 = "no imputation", 1 = "imputed income statement" and -
2 = "does not apply, not working".

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

LABNET\$\$ -- Current net labor income (generated) in euros

Var Label : LABNET\$\$ **Current net labor income (generated) in euros**

Var format : LABNET\$\$ (F2)

19\$\$ - year : \$\$=84..13

Comment:

The variable LABNET\$\$ represents the generated and imputed current net labor income of all persons in SOEP working in the respective wave. Income details are consistently provided in euros for all waves. The imputation of item nonresponse takes place in a two-stage procedure: first, with the "Row-and-Column" method of Little und Su (1989) using individual longitudinal data as well as cross-sectional trend data (cf. Joachim R. Frick and Markus M. Grabka (2005): Item-Non-Response on Income Questions in Panel surveys: Incidence, Imputation and the Impact on the Income Distribution. Allgemeines Statistisches Archiv (ASTA) 89, 49-61). Alternatively, if no individual longitudinal information is available, we base the imputation on a regression using different Mincer covariates, also taking into account current gross labor income. If both types of income information are lacking, first we impute current gross labor income and then current net labor income. Imputed values are flagged(IMPNET\$\$).

The original variables coming from the \$P-files and are:
ap3302,bp4302,cp5202,dp4402,ep4402,fp4502,gp4302,hp5402,
ip5402,jp5402,kp6402,lp5302,mp4702,np5402,op4502,pp6002,
qp5602,rp5702,sp5802,tp7602,up5902,vp7102,wp5902,xp7302,
yp6802,zp7202,bap6102

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

IMPNET\$\$ -- Impututation flag for current net labor income
(LABNET\$\$)

Var Label : IMPNET\$\$ **Impututation flag for current net labor income**
(LABNET\$\$)

Var format : IMPNET\$\$ (I1)

19\$\$ - year : \$\$=84..13

Comment:

The variable IMPNET\$\$ designates imputations of item-
nonresponse in the variable LABNET\$\$ (current net labor
income). IMPNET\$\$ can take the value
0 = "no imputation", 1 = "imputed income statement" and -
2 = "does not apply, not working".

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

SNDJOB\$\$ -- Current gross labor income in euros from second job (generated)

Var Label : SNDJOB\$\$ **Current gross labor income in euros from second job(generated)**

Var format : SNDJOB\$\$ (F2)

19\$\$ - year : \$\$=84..13

Comment: The variable SNDJOB\$\$ represents the imputed current gross labor income from second job generated for all SOEP respondents who are employed 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).

Imputed values are flagged (IMPSND\$\$).

The original variables coming from the \$P-files and are:
lp7702,mp5802,np5802,op4902,pp6602,qp6302,rp6602,sp6602,tp8402,
up67,vp79,wp70,xp81,yp76,zp78,bap69,bbp79,bcp67

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

IMPSND\$\$ -- Imputation flag for SNDJOB\$\$

Var Label : IMPSND\$\$ **Imputation flag for SNDJOB\$\$**

Var format : IMPSND\$\$ (I1)

19\$\$ - year : \$\$=84..13

Comment:

The variable IMPSND\$\$ designates imputations of item-nonresponse in the variable SNDJOB\$\$ (current gross labor income from second job). IMPSND\$\$ can take the value 0 = "no imputation", 1 = "imputed income statement" and - 2 = "does not apply, not working".

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

FIELD\$\$ -- Field of tertiary education

Var Label : FIELD\$\$ "Field of tertiary education"

Value Label : 1-98 (see below)

Var Format : FIELD\$\$ (I2)

19\$\$-Year : \$\$=85..13

[1] Sprach- und Kulturwissenschaften allgemein	[39] Physik, Astronomie
[2] Evang. Theologie, -Religionslehre	[40] Chemie
[3] Kath. Theologie, -Religionslehre	[41] Pharmazie
[4] Philosophie	[42] Biologie
[5] Geschichte	[43] Geowissenschaften (ohne Geographie)
[6] Bibliothekswissenschaft, Dokumentation, Publizistik	[44] Geographie
[7] Allgemeine und vergleichende Literatur und Sprachwissenschaft	[48] Gesundheitswissenschaften allgemein
[8] Altphilologie (klass. Philologie), Neugriechisch	[49] Humanmedizin (ohne Zahnmedizin)
[9] Germanistik (Deutsch, germanische Sprachen ohne Anglistik)	[50] Zahnmedizin
[10] Anglistik, Amerikanistik	[51] Veterinaermedizin
[11] Romanistik	[57] Landespflege, Umweltgestaltung
[12] Slawistik, Baltistik, Finno- Ugristik	[58] Agrarwissenschaften, Lebensmittel- und Getraenketeknologie
[13] Aussereuropaeische Sprach- und Kulturwissenschaften	[59] Forstwissenschaft, Holzwirtschaft
[14] Kulturwissenschaften i.e.S.	[60] Ernaehrungs- und Haushaltswissenschaften
[15] Psychologie	[61] Ingenieurwesen allgemein
[16] Erziehungswissenschaften	[62] Bergbau, Huettenwesen
[17] Sonderpaedagogik	[63] Maschinenbau/Verfahrenstechnik
[22] Sport, Sportwissenschaft	[64] Elektrotechnik
[23] Rechts-, Wirtschafts- und Sozialwissenschaften allgemein	[65] Verkehrstechnik, Nautik
[24] Regionalwissenschaften	[66] Architektur, Innenarchitektur
[25] Politikwissenschaften	[67] Raumplanung
[26] Sozialwissenschaften	[68] Bauingenieurwesen
[27] Sozialwesen	[69] Vermessungswesen
[28] Rechtswissenschaften	[74] Kunst, Kunstwissenschaft allgemein
[29] Verwaltungswissenschaften	[75] Bildende Kunst
[30] Wirtschaftswissenschaften	[76] Gestaltung
[31] Wirtschaftsingenieurwesen	[77] Darstellende Kunst, Film und Fernsehen, Theaterwissenschaft
[36] Mathematik, Naturwissenschaften allgemein	[78] Musik, Musikwissenschaft
[37] Mathematik	[83] Ausserhalb der Studienbereichsgliederung
[38] Informatik	[98] nicht zuzuordnen

Comment: The variable is designed to provide information on the field of education of tertiary degrees which adds details to the information recorded in the variable \$PBBIL02. While the latter variable records if a person holds a degree FIELD\$\$ contains more detailed information on the type of the degree. The data

Documentation of the wave-specific files \$PGEN

of the generated variable FIELD\$\$ stem from two sources: 1. Person questionnaire: Each year since 1985 respondents are asked if they have left education since the beginning of the year prior to the survey and which degrees they have obtained. This part of the questionnaire contains an open question on the type and the field of newly obtained tertiary degrees. This information is coded and used for the generation of the variables FIELD\$\$.

2. Biography questionnaire: Since 2001 similar information is collected from respondents who fill in the biography questionnaire (usually during the first two years of participation in the panel). In contrast to the information from the person questionnaire the questions do not refer to currently obtained degrees but to degrees obtained during the time before being part of the SOEP sample.

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

Each year the variable contains the most recently collected information. Take for instance a person for whom we have observed a first degree in sociology in 1987 and a second degree in economics in 1991. For this person the variables FIELD\$\$ would be filled as follows:

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

If you want to take into account that a person holds two degrees you have to combine the information from all available years. However, only a minority of the population holds more than one tertiary degree. In very few cases we encounter the problem that a respondent provides information on two different degrees in one survey year. This only happens in years when respondents fill in the person as well as the biography questionnaire. In these cases we prioritize the information from the person questionnaire as it refers to the current situation while the biography questionnaire contains retrospective information. Furthermore, there are cases who report an applied university degree and a university degree in the biography questionnaire. In these cases, the variable contains information on the university degree only.

The variable is coded according to the classification on fields of education ("Fächergruppen") provided by the Statistisches Bundesamt (2009). Until 2009 data from the person questionnaire were coded using an earlier version of this classification (1982). In the variable FIELD\$\$ we recoded the original values. As the newer version is more precise this could be done with hardly any loss of information. Some categories are collapsed. Category 3 is coded as 2 (no distinction between catholic and protestant theology), 14 as 13, 17 as 16, 24 as 23, 25 as 26 and 48 as 49. The original values of the data collected from the person questionnaire up to 2009 are stored in the respective variables in the dataset \$P.

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

Documentation of the wave-specific files \$PGEN

Detailed description: Stat. Bundesamt (2009): Bildung und Kultur. Studierende an Hochschulen, Fachserie 11 Reihe 4.1, Wiesbaden: 446ff, Übersicht 1: „Fächergruppen, Studienbereiche und Studienfächer“.

DEGREE\$\$ -- Type of tertiary degree

Var Label : DEGREE\$\$ "Type of tertiary degree"
Value Label : 11-98 (see below)

Var Format : DEGREE\$\$ (I2)
19\$\$-Year : \$\$=85..13

- [11] Magister
- [12] Diplom (Universitaet)
- [13] Bachelor, nicht Lehramt (Universitaet)
- [14] Master, nicht Lehramt (Universitaet)
- [15] Erstes Staatsexamen, nicht Lehramt
- [16] sonstige Staatsexamen, nicht Lehramt
- [21] Diplom (FH, Verwaltungsfachhochschule)
- [22] Bachelor, nicht Lehramt (FH, Verwaltungsfachhochschule)
- [23] Master, nicht Lehramt (FH, Verwaltungsfachhochschule)
- [31] Lehramt,Bachelor,Master an Grund-,Hauptschulen/Primarstufe
- [32] Lehramt,Bachelor,Master, Sekundarstufe I/Grundschulen/Primarstufe
- [33] Lehramt,Bachelor,Master, an Realschulen/Sekundarstufe I
- [34] Lehramt,Bachelor,Master, Sekundarstufe II und I
- [35] Lehramt,Bachelor,Master, Gymnasien/Sekundarstufe II, allg.bil.Schulen
- [36] Lehramt,Bachelor,Master, an Sonder-/Foerderschulen
- [37] Lehramt,Bachelor,Master, an beruflichen Schulen
- [38] Lehramt,sonstiges
- [41] kuenstlerischer Abschluss
- [42] Promotion
- [43] Habilitation
- [44] Sonstiger Abschluss
- [98] nicht zuzuordnen

Comment: The variable is designed to provide information on the type of tertiary degree (e.g., Diploma, Bachelor, Master) which adds details to the information recorded in the variable \$PBBIL02. While the latter variable records if a persons holds a degree DEGREE\$\$ contains more detailed information on the type of the degree. The data of the generated variable DEGREE\$\$ stem from two sources: 1. Person questionnaire: Each year since 1985 respondents are asked if they have left education since the beginning of the year prior to the survey and which degrees they have obtained. This part of the questionnaire contains an open question on the type and the field of newly obtained tertiary degrees. This information is coded and used for the generation of the variables DEGREE\$\$.

2. Biography questionnaire: Since 2001 similar information is collected from respondents who fill in the biography questionnaire (usually during the first two years of participation in the panel). In contrast to the information from the person questionnaire the questions do not refer to currently obtained degrees but to degrees obtained during the time before being part of the SOEP sample.

In the variable DEGREE\$\$ we combine these two types of information. However, since the retrospective information was not collected before 2001 the variable covers until 2000 only persons for whom we have prospectively observed the end of study. This explains why the number of valid observations is

Documentation of the wave-specific files \$PGEN

rather small in these years. Information on the data source is stored in the variable FDT_F\$\$.

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

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

If you want to take into account that a person holds two degrees you have to combine the information from all available years. However, only a minority of the population holds more than one tertiary degree. In very few cases we encounter the problem that a respondent provides information on two different degrees in one survey year. This only happens in years when respondents fill in the person as well as the biography questionnaire. In these cases we prioritize the information from the person questionnaire as it refers to the current situation while the biography questionnaire contains retrospective information. Furthermore, there are cases who report an applied university degree and a university degree in the biography questionnaire. In these cases, the variables contain information on the university degree only.

The variable is coded according to a slightly collapsed version of the classification on types of tertiary degrees ("Prüfungsgruppen und Abschlussprüfungen") provided by the Statistisches Bundesamt (2009). Until 2009 data from the person questionnaire were coded using an earlier version of this classification (1982) which was slightly revised in 2009 (inclusion of Bachelor and Master degrees). Since 2010 the data were coded according to the classification presented here. In the variable DEGREE\$\$ we recoded the original values from years 2009 and earlier. As the newer version is more precise this could be done with hardly any loss of information. Some categories are collapsed. Category 16 was mostly likely coded as 15 in earlier years, 34 as 35 and 43 as 44. The original values of the data collected from the person questionnaire up to 2009 are stored in the respective variables in the dataset \$P.

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

Detailed description: Stat. Bundesamt (2009): Bildung und Kultur. Studierende an Hochschulen, Fachserie 11 Reihe 4.1, Wiesbaden: 449ff, Übersicht 2: „Prüfungsgruppen und Abschlussprüfungen“.

TRAINA\$\$ -- Apprenticeship - two-digit occupation KldB92

Var Label: TRAINA\$\$ "Apprenticeship - two-digit occupation KldB92"

Value Label : 1-99 (see below)

Var Format : TRAINA\$\$ (I2)
19\$\$-Year : \$\$=85..13

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

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

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

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

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

If you want to take into account that a person holds two vocational qualifications you have to combine the information from all available years. In few cases we encounter the problem

that a respondent provides information on two different apprenticeships in one survey year. This only happens once, namely in years when respondents fill in the person as well as the biography questionnaire. In these cases we prioritize the information from the person questionnaire as it refers to the current situation while the biography questionnaire contains retrospective information.

The variable is coded according to the classification of occupations at two-digit level ("Berufsgruppen") provided by the Statistisches Bundesamt (1992). Other SOEP occupation variables are coded at four-digit level. The reason why the variable TRAINA\$\$ is provided at two-digit level only is that until 2009 the data from the two different sources were coded according two different classifications which could be combined at a higher level of aggregation only. The person questionnaire data were coded according to the classification of occupations provided by the Bundesanstalt für Arbeit (1988, four-digit level) while the biography data use the classification provided by the Statistisches Bundesamt (1992, four-digit level). Since 2010 both types of data are coded according to the latter classification. The four-digit version of the biography data can be provided upon request. The original values of the data collected from the person questionnaire up to 2009 are stored in the respective variables in the dataset \$P.

Please note that for respondents from the newly introduced migration sample a valid value to TRAINA\$\$ was only assigned if the vocational training was completed in Germany.

Detailed description: Hartmann/Schütz (2002): Die Klassifikation der Berufe und der Wirtschaftszweige im Sozio-oekonomischen Panel - Neuvercodung der Daten 1984 - 2001. Infratest Sozialforschung, München.

TRAINB\$\$ -- Vocational school - two-digit occupation KldB92

Var Label: TRAINB\$\$ "Vocational school - two-digit occupation KldB92"

Value Label : 1-99 (see TRAINA\$\$)

Var Format: TRAINB\$\$ (I2)
19\$\$-Year : \$\$=85..13

Comment: The variable is designed to provide information on the occupation of full-time school based vocational training (e.g., Berufsfachschule, Schule des Gesundheitswesens, Handelsschule). See the description of variable TRAINA\$\$ for more details on the construction and the values of the variable.

TRAINC\$\$ -- Higher vocational school - two-digit occupation KldB92

Var Label : TRAINC\$\$ "Higher vocational school - two-digit occupation KldB92"

Value Label : 1-99 (see TRAINA\$\$)

Var Format : TRAINC\$\$ (I2)
19\$\$-Year : \$\$=85..13

Comment: The variable is designed to provide information on the occupation of higher level vocational training (e.g., Meister, Techniker). See the description of variable TRAINA\$\$ for more details on the construction and the values of the variable.

TRAIND\$\$ --Civil servant training - two-digit occupation KldB92

Var Label : TRAIND\$\$ "Civil servant training - two-digit occupation KldB92"

Value Label : 1-99 (see TRAINA\$\$)

Var Format : TRAIND\$\$ (I2)
19\$\$-Year : \$\$=85..13

Comment: The variable is designed to provide information on the occupation of civil servant training ("Beamtenausbildung"). See the description of variable TRAINA\$\$ for more details on the construction and the values of the variable.

FDT_F\$\$ -- Data source FIELD, DEGREE, TRAIN

Var Label : FDT_F\$\$ "Data source FIELD, DEGREE, TRAIN"

Value Label : (1) person questionnaire
(2) person questionnaire (temporary drop-out)
(3) biography questionnaire
(4) various sources

Var Format : FDT_F\$\$ (I1)
19\$\$-Year : \$\$=85..13

Comment: This is a flag variable which provides information on the data sources used for the construction of the variables FIELD\$\$, DEGREE\$\$, TRAINA\$\$, TRAINB\$\$, TRAINC\$\$ and TRAIND\$\$ (see the description of the respective variables for details).

BILZTCH\$\$

Var Label : BILZTCH\$\$ **Change in highest educational attainment since last observation/year**

Value Label : BILZTCH\$\$
(0) 'consistent educational information since last observation'
(1) 'inconsistent educational decrease since last observation'
(2) 'inconsistent educational increase since last year'

Var format : BILZTCH\$\$ (I1)
19\$\$-Year : \$\$=85..13

Comment: 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, all SOEP participants were asked again to state their highest level of educational attainment. This produced a series of inconsistencies between the most recent information from 2000 and the generated information from previous years, which had been based in part on information collected many years prior. These inconsistencies in response behaviour, which include both upward and downward shifts in the values for highest educational attainment, 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, making it easier for users to deal appropriately with such inconsistencies. 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. Since 2011, a Beta version of BIOEDU has also been made available, containing new data on consistent longitudinally tested educational transitions.

For more information, contact: Pia Schober (Tel. +49-30-89789-691)

BILZTEV\$\$

Var Label : BILZTEV\$\$ **Change in highest educational attainment over the whole observation period**

Value Label : BILZTEV\$\$
(0) 'consistent educational information'
(1) 'inconsistent decrease in educational qualifications'
(2) 'inconsistent increase in educational qualifications'
(3) 'inconsistent increase and decrease in qualifications'

Var format : BILZTEV\$\$ (I1)
19\$\$-Year : \$\$=85..13

Comment: This is a flag variable which identifies observations which ever had one or more inconsistent changes in the information given on their highest educational qualification over the whole observation period. See the description of variable BILZTCH\$\$ for more details on the sources of these inconsistencies.

For more information, contact: Pia Schober (Tel. +49-30-89789-691)

Variables no longer distributed:

ISCOU\$\$

This variable based on ISCO68 has been discontinued since 2000. Instead, all past answers provided by respondents in their own words were recoded based on the ISCO88 procedure; this hierarchically structured occupational classification scheme replaces the previous differentiation into 1-digit, 2-digit and 3-digit occupational codes.

The codes of this variable generated since 2000 can be obtained upon request from the SOEP group. Since 2001 the variables from past waves have been erased.

ISCO\$\$

This variable based on ISCO68 has been discontinued since 2000. Instead all past answers provided by respondents in their own words were recoded based on the ISCO88 procedure; this hierarchically structured occupational classification scheme replaces the previous differentiation into 1-digit, 2-digit and 3-digit occupational codes.

The codes of this variable generated since 2000 can be obtained upon request from the SOEP group. Since 2001 the variables from past waves have been erased.

ISCOH\$\$

This variable based on ISCO68 has been discontinued since 2000. Instead all past answers provided by respondents in their own words were recoded based on the ISCO88 procedure; this hierarchically structured occupational classification scheme replaces the previous differentiation into 1-digit, 2-digit and 3-digit occupational codes.

The codes of this variable generated since 2000 can be obtained upon request from the SOEP group. Since 2001 the variables from past waves have been erased.

BRANCH\$\$

This variable is the result of the answers provided by respondents in their own words regarding economic sectors. These answers, which cannot be passed on to data users due to data protection regulations, were recoded following a survey carried out by the ZUMA Center for Survey Research and Methodology in Mannheim, according to an expanded sector list developed by DIW Berlin. This list was based on ZUMA “standard demographics” list, which is completely comparable up to Code 23. Since the classification based on NACE enables international comparison, no codes were assigned based on the earlier sectoral classification.

The codes of this variable generated up to 2000 can be obtained on request from the SOEP group. Since 2001 the variables from past waves have been erased.