Research Report

SOEP 2011 - Documentation of person-related status and generated variables in PGEN for SOEP.v28

SOEP Survey Papers, No. 111

Provided in Cooperation with:
German Institute for Economic Research (DIW Berlin)

Suggested Citation: DIW Berlin / SOEP (Ed.) (2012): SOEP 2011 - Documentation of person-related status and generated variables in PGEN for SOEP.v28, SOEP Survey Papers, No. 111, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:
http://hdl.handle.net/10419/68151

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
SOEP 2011 – Documentation of Person-related Status and Generated Variables in PGEN for SOEP.v28

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

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

The SOEP Survey Papers is comprised of the following series:

Series A – Survey Instruments (Erhebungsinstrumente)
Series B – Survey Reports (Methodenberichte)
Series C – Data Documentation (Datendokumentationen)
Series D – Variable Descriptions and Coding
Series E – SOEPtrackers
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:

ISSN: 2193-5580 (online)

Contact: DIW Berlin
SOEP
Mohrenstr. 58
10117 Berlin

Email: soeppapers@diw.de
Documentation PGEN

Person-related status and generated variables
List of variables:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERWTYPSS$</td>
<td>3</td>
</tr>
<tr>
<td>JOBCHSS$</td>
<td>4</td>
</tr>
<tr>
<td>EMPLSTSS$</td>
<td>5</td>
</tr>
<tr>
<td>LFSSS$</td>
<td>6</td>
</tr>
<tr>
<td>EXPFTSS$</td>
<td>7</td>
</tr>
<tr>
<td>EXPPTSS$</td>
<td>8</td>
</tr>
<tr>
<td>EXPUEESS$</td>
<td>9</td>
</tr>
<tr>
<td>SERWZEIT$</td>
<td>10</td>
</tr>
<tr>
<td>ERLJOBSS$</td>
<td>12</td>
</tr>
<tr>
<td>AUSBSS$</td>
<td>13</td>
</tr>
<tr>
<td>BETRSS$</td>
<td>15</td>
</tr>
<tr>
<td>ALLBETSS$</td>
<td>17</td>
</tr>
<tr>
<td>OEFFDSS$</td>
<td>18</td>
</tr>
<tr>
<td>STIBSS$</td>
<td>19</td>
</tr>
<tr>
<td>KLAASS$</td>
<td>21</td>
</tr>
<tr>
<td>IS88SS$</td>
<td>23</td>
</tr>
<tr>
<td>NACESS$</td>
<td>25</td>
</tr>
<tr>
<td>ISEISS$</td>
<td>27</td>
</tr>
<tr>
<td>EGPSS$</td>
<td>28</td>
</tr>
<tr>
<td>SIOPSS$</td>
<td>30</td>
</tr>
<tr>
<td>MPSSS$</td>
<td>31</td>
</tr>
<tr>
<td>AUTONOSS$</td>
<td>32</td>
</tr>
<tr>
<td>SVEBZEIT$</td>
<td>33</td>
</tr>
<tr>
<td>STATZEIT$</td>
<td>34</td>
</tr>
<tr>
<td>SUEBSTD$</td>
<td>35</td>
</tr>
<tr>
<td>PARTZSS$</td>
<td>36</td>
</tr>
<tr>
<td>PARTNRS$</td>
<td>37</td>
</tr>
<tr>
<td>SFAMSTD$</td>
<td>38</td>
</tr>
<tr>
<td>NATIONSS$</td>
<td>39</td>
</tr>
<tr>
<td>SPSBIL</td>
<td>40</td>
</tr>
<tr>
<td>SPSBIL0</td>
<td>46</td>
</tr>
<tr>
<td>SPSBILA</td>
<td>47</td>
</tr>
<tr>
<td>SPBBIL01</td>
<td>48</td>
</tr>
<tr>
<td>SPBBIL0</td>
<td>49</td>
</tr>
<tr>
<td>SPBBIL02</td>
<td>50</td>
</tr>
<tr>
<td>SPBBIL03</td>
<td>51</td>
</tr>
<tr>
<td>SPBBILA</td>
<td>52</td>
</tr>
<tr>
<td>SBILZEIT</td>
<td>53</td>
</tr>
<tr>
<td>SISCED</td>
<td>55</td>
</tr>
<tr>
<td>SCASMIN</td>
<td>57</td>
</tr>
<tr>
<td>MONTHSS$</td>
<td>59</td>
</tr>
<tr>
<td>MODESS$</td>
<td>60</td>
</tr>
<tr>
<td>LABGROSS$</td>
<td>61</td>
</tr>
<tr>
<td>IMPGRSS$</td>
<td>62</td>
</tr>
<tr>
<td>LABNETS$</td>
<td>63</td>
</tr>
<tr>
<td>IMPNETSS$</td>
<td>64</td>
</tr>
<tr>
<td>FIELDSS$</td>
<td>65</td>
</tr>
<tr>
<td>DEGREESS$</td>
<td>67</td>
</tr>
<tr>
<td>TRAINASS$</td>
<td>69</td>
</tr>
<tr>
<td>TRAINBSS$</td>
<td>71</td>
</tr>
<tr>
<td>TRAINCSS$</td>
<td>71</td>
</tr>
<tr>
<td>TRAININDSS$</td>
<td>71</td>
</tr>
</tbody>
</table>

Variables no longer distributed:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISCOUSS$</td>
<td>73</td>
</tr>
<tr>
<td>ISCOSS$</td>
<td>73</td>
</tr>
<tr>
<td>ISCOHSS$</td>
<td>73</td>
</tr>
<tr>
<td>BRANCHSS$</td>
<td>73</td>
</tr>
</tbody>
</table>
ERWTYP$$

Var Label :  ERWTYP$$  "Type of occupation"
Value Label :  ERWTYP$$  (-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..11

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 and 2010 — 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 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:  Silke Anger (Tel. +49-30-89789-526)
**JOBCHE$$**

Var Label : JOBCHE$$  "Job change"
Value Label : JOBCHE$$ (-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 : JOCHCHE$$ (I2)
19$$ - Year : $$=84..11

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.

JOCHCHE$$ 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, JOBCHE$$ is a modified version of the variable ERWTYP$$, 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$$, the variable is designed to provide consistent longitudinal information on job changes. The JOBCHE$$ variable is generated by correcting the original job change information in various ways:
1. We check whether the job changes stated by a respondent in two consecutives interviews refer to one and the same job change. The date of the job change and the interview month are used to correct double entries.
2. If the respondent indicates a job change with a date before the previous interview but did not state a job change in the previous interview, this is coded as a job change in the current interview.
3. If a respondent indicates no job change and was not employed at the time of the previous interview, this is coded as "no job change" despite the seeming implausibility, since there are possible explanations how this information could be plausible, e.g. if there were short-term employment spells between two interview dates.
4. Respondents can be "first-time employed" only once. If a respondent states being "first-time employed" for a second time, this is coded as "employed, with change".

For more information, contact: Silke Anger (Tel. +49-30-89789-526)
Var Label : EMPLST$$ “Employment status”
Value Labels : EMPLST$$
(-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 : EMPLST$$ (I2)
19$$ - year : $$=84..97

Var Label : EMPLST$$ “Employment status”
Value Labels : EMPLST$$
(-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 : EMPLST$$ (I2)
19$$ - year : $$=98..11

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$$ supplements the variable LFS$$, which differentiates among persons who are not employed.

For more information, contact: Silke Anger (Tel. +49-30-89789-526)
**LFS$$**

Var Label : LFS$$  "Labor force status"
Value Label : LFS$$ (-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 : LFS$$ (I2)
19$$ - Year : $$=84..11

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.

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

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

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

LFS$$ supplements the variable EMPLST$$, which differentiates among persons who are employed.

For more information, contact:  Silke Anger (Tel. +49-30-89789-526)
**EXPFT$$**

Var Label : EXPFT$$ “Working experience, full-time employment”

Value Labels : EXPFT$$ (-3) “implausible answer”
               (-1) “no answer”

Var format : EXPFT$$ (I3)
19$$ - year : $$=84..11

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

If there is no monthly calendar data available in a given year of a respondent’s career, the annual data from PBIOSPE is used for that year. In the most current wave the variable only uses up-to-date information from the newly answered Biography Questionnaires. If the year in which a spell started and ended is the same, and if there is no monthly data, a spell of 0.5 years is assumed. Persons 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: Silke Anger (Tel. +49-30-89789-526)
**EXPPT**

Var Label: **EXPPT** “Working experience, part-time employment”

Value Labels: **EXPPT**
   (-3) “implausible answer”
   (-1) “no answer”

Var format: **EXPPT** (I3)

1999 - year: $$=84..11

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

If there is no monthly calendar data available in a given year of a respondent’s career, the annual data from **PBIOSPE** is used for that year. In the most current wave the variable only uses up-to-date information from the newly answered Biography Questionnaires. If the year in which a spell started and ended is the same, and if there is no monthly data, a spell of 0.5 years is assumed. Persons 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: Silke Anger (Tel. +49-30-89789-526)
**EXPUE**

Var Label : EXPUE$\$ "Unemployment experience"
Value Labels : EXPUE$\$ (-3)"implausible answer"
               (-1)"no answer"
Var format : EXPUE$\$ (I3)
19$$ - year : $$=84..11

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

If there is no monthly calendar data available on a given year in a respondent’s career, the annual data from PBIOSPE is used for that year. In the most current wave the variable only uses up-to-date information from the newly answered Biography Questionnaires. If the year in which a spell started and ended is the same, and if there is no monthly data, a spell of 0.5 years is assumed. Persons 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$$ and EXPFT $$.

For more information, contact: Silke Anger (Tel. +49-30-89789-526)
$ERWZEIT

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..11

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: Silke Anger (Tel. +49-30-89789-526)
**ERLJOB$$**

**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..11

**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 and 2010 — 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: Silke Anger (Tel. +49-30-89789-526)
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$$ (I2)
19$$ - Jahr : $$=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..11

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 and 2010 - 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: Silke Anger (Tel. +49-30-89789-526)
**BETR$\$**

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”
              (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 : $$=91.98

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$\$ (I2)
+19$\$ - 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 : $$=05.11

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 - 1985, 1986, 1987, 1988, 1990 (West), 1992 (West), 1994, 1996, 1999, 2001, 2003, 2005, 2006, 2008 and 2010 - 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.

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: Silke Anger (Tel. +49-30-89789-526)
**ALLBET$$**

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..11

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 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.

For more information, contact: Silke Anger (Tel. +49-30-89789-526)
OEFFD$$

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..11

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 and 2010 - 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: Silke Anger (Tel. +49-30-89789-526)
**STIB$$**

Var Label : stib$$
Value Labels : stib$$ (1-640)(see below)
Var format : STIB$$ (I3)

19$$ - Year : $$=84..11

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

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: Silke Anger (Tel. +49-30-89789-526)
KLASS$

Var Label : KLASS$  "StaBuA 1992 Job Classification"
Value Label : KLASS$  (-3)"no information for KLAS"
               (-1)"no answer"
               (-2)"not applicable - not employed"

Var format : KLASS$
19$$ - Year : $$=84..11

Comment: This variable is designed to provide annual data on job classification for all employed persons according to the classification of the German Federal Statistical Office (StaBuA). Respondents answer the question on their current occupational title in their own words, and this response is entered into a blank in the questionnaire. Due to data protection regulations, this information cannot be provided to data users and was therefore completely recoded by Infratest Sozialforschung in the year 2002. This recoding has been documented in Hartmann/Schütz 2002.

Not all employed persons are asked 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, 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, KLASS$ 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.

The occupational classification of the German Federal Statistical Office differentiates among six main occupational types (see next page):
For more information, contact: Marco Giesselmann (Tel. +49-30-89789-503)

Because of gaps in the answers provided by respondents, the following “new” codes were created:

9711 - Mithelfende Familienangehörige außerhalb der Landwirtschaft, anderweitig nicht genannt
9811 - Auszubildende mit (noch) nicht feststehendem Ausbildungsbereif
9821 - Praktikanten/Praktikantinnen, Volontäre/ Volontärianne mit (noch) nicht feststehendem Beruf
9911 - Facharbeiter/innen, ohne nähere Tätigkeitsangabe
9921 - Heinarbeiter/innen, ohne nähere Tätigkeitsangabe
9931 - Vorarbeiter/innen, Gruppenleiter/innen, ohne nähere Tätigkeitsangabe
9971 - Sonstige Arbeitskräfte, ohne nähere Tätigkeitsangabe


Var Label : ISCO$$ “4-digit ISCO-88 Occupation Code”
Value Label : ISCO$$ (-3)”no information for ISCO”
( -1) ”no answer”
(-2) ”not applicable – not employed”

Code name (Main group, group):

(1000) Legislators, senior officials, and managers
(1100) Legislators and senior officials
(1200) Corporate managers
(1300) Managers of small enterprises
(2000) Professionals
(2100) Physical, mathematical, and engineering science professionals
(2200) Life science and health professionals
(2300) Teaching professionals
(2400) Other professionals
(3000) Technicians and associate professionals
(3100) Physical and engineering science associate professionals
(3200) Life science and health associate professionals
(3300) Teaching associate professionals
(3400) Other associate professionals
(4000) Clerks
(4100) Office clerks
(4200) Customer services clerks
(5000) Service Workers and shop and market sales workers
(5100) Personal and protective services workers
(5200) Models, salespersons, and demonstrators
(6000) Skilled agricultural and fishery workers
(6100) Skilled agricultural and fishery workers
(7000) Craft and related trades workers
(7100) Extraction and building trades workers
(7200) Metal, machinery, and related trades workers
(7300) Precision, handicraft, craft, printing and related trades workers
(7400) Other craft and related trades workers
(8000) Plant and machine operators and assemblers
(8100) Stationary plant and related operators
(8200) Machine operators and assemblers
(8300) Drivers and mobile plant operators
(9000) Elementary occupations
(9100) Sales and services elementary occupations
(9200) Agricultural, fishery, and related laborers
(9300) Laborers in mining, construction, manufacturing, and transport

Var format : IS88$$ (I4)
19$$ - year : $$=84 - 11

Comment: This variable is designed to provide annual data on occupational activity for all employed persons according to the International Standard Classification of Occupations ISCO-88. Respondents answer the question on their current occupational title in their own words, and this response is entered into a blank in the questionnaire. Due to data protection regulations, this information cannot be provided to data users and was therefore completely recoded by Infratest Sozialforschung in the year 2002. This recoding has been documented in Hartmann/Schütz 2002.

ISCO-88 is a strictly four-digit classification, and this variable is therefore coded in four-digit form. In contrast to the previous version of the classification system, ISCO-68, ISCO-88 does not use blanks if there is not adequate information for specific coding, but uses zeros instead. Thus 4000 stands for an unspecified office job; 2300 stands for teachers and 2000 stands for scientists, both without closer specification. There is no conversion key since the two classifications differ significantly. The SOEP data distribution 1984-2001 replaced ALL earlier data distributions with the ISCO-88-coding.

up-to-date information. Therefore, in years with a partial survey, IS88$$ usually contains available information from the previous year for persons without a job change (persons with JOBCH$$-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.

**Detailed description:**


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

Var Label : NACE$$  "Two-digit NACE Industry - Sector"
Value Label : NACE$$ (1-100) (see below)
Var format : NACE$$ (I2)
19$$ - Year : $$=84..11

( -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

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 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.

SOEP v28
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 and 2010 - 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.


For more information, contact: Silke Anger (Tel. +49-30-89789-526)
**ISEI**

**Var Label**: ISEI

**Value Label**: ISEI

- (-3) "No information for ISEI"
- (-1) "No answer"
- (-2) "Not applicable - not employed"

**Var format**: ISEI

**19** - Year: = 84 - 11

**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 (SIOPS, MPS$) and occupational class (EGP$).


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

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 - 11

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$$) and occupational prestige scores (SIOPSS$$, MPSS$$).


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

Var Label : SIOPSS$
Value Label : SIOPSS$
(-3) "no information for SIOPS"
(-2) "not employed"
(-1) "no answer"

Var format : SIOPSS$ (I2)
19$$ - Year : $$=84..11

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$$).


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

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..11

Comment: This variable gives the occupational prestige score developed by Wegener (1988) for all employed persons. Like the SIOPSS$$ prestige sore, Wegener’s prestige scala measures a person’s occupational prestige and was developed especially for use in the Federal Republic of Germany. MPS$$ is assigned based on the German Federal Statistical Office’s occupational classification of 1992 (KLASS$$. 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 (SIOPSS$$), occupational status (ISEI$$), and occupational class (EGP$$).


For more information, contact: Marco Giesselmann (Tel. +49-30-89789-503)
**Var Label**: AUTONO$$

**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..11

**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 (SIOPS$$). AUTONO$$ is the simplest variable based on the scales of "occupational position" in terms of its construction, and strongly correlated with the Treiman Prestige Scale (SIOPS$$).

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

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


**For more information, contact**: Silke Anger (Tel. +49-30-89789-526)
$VEBZEIT

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..11

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 $TATZEIT and $UEBSTD.

For more information, contact: Silke Anger (Tel. +49-30-89789-526)
**$TATZEIT**

Var Label : $TATZEIT “Actual weekly working hours”

Value Labels : $TATZEIT (-3)"implausible answer"
(-2)"does not apply - not employed"
(-1)"no answer"

Var format : $TATZEIT (I3)

$\$ - Year : $\$=84..11

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: Silke Anger (Tel. +49-30-89789-526)
**$UEBSTD**

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..11  (not in 1987)

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 for 1984 and 1985 were therefore generated using the difference between average actual weekly working hours and agreed weekly working hours. It is not possible to create this kind of variable for the year 1987, since all values here were "missings". 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: Silke Anger (Tel. +49-30-89789-526)
**PARTZ$$**

**Var Label**: PARTZ$$ Partner indicator
**Value Label**: PARTZ$$
- (0) no partner, clearly
- (1) spouse, clearly
- (2) partner, clearly
- (3) probably spouse
- (4) probably partner
- (9) partner exists, identity unknown

**Var format**: PARTZ$$ (I1)
**19$$ - year : $$=84..11

**COMMENT**: Partner indicators have the purpose of clearly defining spouse (married) and partner (unmarried) relationships 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 automatically assigned to all persons living in households in which there is clearly no partnership. These include:
  - (a) one-person households
  - (b) single-parent households
  - (c) household head living together with only one parent (or parent-in-law)
- Codes 1 to 4 define these relationships. To assign Codes 1 and 2, the partnership has to be clearly definable from the perspective of both partners. This implies agreement between the codes of the variable $STELL (= relationship to head of household in $PBRUTTO) pointing to a possible partnership (e.g., the combination 0 (=head of household) and 1 (=spouse of household head)), as well as agreement between the codes for family status in that wave (e.g., married couples both have the Code 1 (=married, living together)). In case of unclarities, the marital history is taken into account as well. If there are inconsistencies between the answers provided by the two persons, or between data on marital status and relationship to head of household, each person is examined individually within his or her household context. If uncertainty remains, the codes 3 or 4 are assigned. Code 9 is assigned if at least two other household members might potentially be a particular person’s partner and thus no clear determination of partnership can be made (usually only the case among households in Sample B).

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

Var Label : PARTNR$$  Person ID number of partner
Var format : PARTNR$$ (I6)
19$$ - year : $$=84..11

**COMMENT:** Partner indicators have the purpose of clearly defining spouse (married) and partner (unmarried) relationships in SOEP households and thus to make possible analyses on the couple level. If PARTZ$$ is coded 0 or 9, this person has no partner or the partner cannot be identified as such. The variable PARTNR$$ 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. For analyses of partner relationships, this information can be used to clearly link all persons with their respective partners, and all information on both partners can also be stored in a common dataset.

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

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)

$5 - year : $5=84..11

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)
<table>
<thead>
<tr>
<th>Country</th>
<th>Value Label</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>( 1) Germany</td>
<td>Germany</td>
</tr>
<tr>
<td>Turkey</td>
<td>( 2) Turkey</td>
<td>( 56) New Zealand</td>
</tr>
<tr>
<td>Ex-Yugoslavia</td>
<td>( 3) Ex-Yugoslavia</td>
<td>( 57) Tanzania</td>
</tr>
<tr>
<td>Greece</td>
<td>( 4) Greece</td>
<td>( 58) Cyprus</td>
</tr>
<tr>
<td>Italy</td>
<td>( 5) Italy</td>
<td>( 59) Cuba</td>
</tr>
<tr>
<td>Spain</td>
<td>( 6) Spain</td>
<td>( 60) Iraq</td>
</tr>
<tr>
<td>Ex-GDR (Only Country Of Origin)</td>
<td>( 7) Ex-GDR (Only Country Of Origin)</td>
<td>( 61) Brazil</td>
</tr>
<tr>
<td>Austria</td>
<td>( 10) Austria</td>
<td>( 62) Monaco</td>
</tr>
<tr>
<td>France</td>
<td>( 11) France</td>
<td>( 63) Hong Kong</td>
</tr>
<tr>
<td>Benelux</td>
<td>( 12) Benelux</td>
<td>( 64) Peru</td>
</tr>
<tr>
<td>Denmark</td>
<td>( 13) Denmark</td>
<td>( 65) Sri Lanka</td>
</tr>
<tr>
<td>Great Britain</td>
<td>( 14) Great Britain</td>
<td>( 66) Nepal</td>
</tr>
<tr>
<td>Sweden</td>
<td>( 15) Sweden</td>
<td>( 67) Morocco</td>
</tr>
<tr>
<td>Norway</td>
<td>( 16) Norway</td>
<td>( 68) China</td>
</tr>
<tr>
<td>Finland</td>
<td>( 17) Finland</td>
<td>( 69) Liechtenstein</td>
</tr>
<tr>
<td>USA</td>
<td>( 18) USA</td>
<td>( 70) Iceland</td>
</tr>
<tr>
<td>Switzerland</td>
<td>( 19) Switzerland</td>
<td>( 71) Ireland</td>
</tr>
<tr>
<td>Chile</td>
<td>( 20) Chile</td>
<td>( 72) St. Lucia</td>
</tr>
<tr>
<td>Romania</td>
<td>( 21) Romania</td>
<td>( 73) Moldavia</td>
</tr>
<tr>
<td>Poland</td>
<td>( 22) Poland</td>
<td>( 74) Kazakhstan</td>
</tr>
<tr>
<td>Korea</td>
<td>( 23) Korea</td>
<td>( 75) Albania</td>
</tr>
<tr>
<td>Iran</td>
<td>( 24) Iran</td>
<td>( 76) Lebanon</td>
</tr>
<tr>
<td>Indonesia</td>
<td>( 25) Indonesia</td>
<td>( 77) Kyrgyzstan</td>
</tr>
<tr>
<td>Hungary</td>
<td>( 26) Hungary</td>
<td>( 78) Ukraine</td>
</tr>
<tr>
<td>Bolivia</td>
<td>( 27) Bolivia</td>
<td>( 79) Algeria</td>
</tr>
<tr>
<td>Portugal</td>
<td>( 28) Portugal</td>
<td>( 80) Mozambique</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>( 29) Bulgaria</td>
<td>( 81) Egypt</td>
</tr>
<tr>
<td>Syria</td>
<td>( 30) Syria</td>
<td>( 81) Egypt</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>( 31) Czech Republic</td>
<td>( 82) Tajikistan</td>
</tr>
<tr>
<td>Russia</td>
<td>( 32) Russia</td>
<td>( 83) Vietnam</td>
</tr>
<tr>
<td>Empty (was Kurdistan)</td>
<td>( 33) Empty (was Kurdistan)</td>
<td>( 84) Somalia</td>
</tr>
<tr>
<td>Mexico</td>
<td>( 34) Mexico</td>
<td>( 85) Pakistan</td>
</tr>
<tr>
<td>Argentina</td>
<td>( 35) Argentina</td>
<td>( 86) South Africa</td>
</tr>
<tr>
<td>Cap Verde Is.</td>
<td>( 36) Cap Verde Is.</td>
<td>( 87) UAE</td>
</tr>
<tr>
<td>Benin</td>
<td>( 37) Benin</td>
<td>( 88) El Salvador</td>
</tr>
<tr>
<td>Philippines</td>
<td>( 38) Philippines</td>
<td>( 89) Eritrea</td>
</tr>
<tr>
<td>Israel</td>
<td>( 39) Israel</td>
<td>( 90) Jordan</td>
</tr>
<tr>
<td>Japan</td>
<td>( 40) Japan</td>
<td>( 91) Turkmenistan</td>
</tr>
<tr>
<td>Australia</td>
<td>( 41) Australia</td>
<td>( 92) Costa Rica</td>
</tr>
<tr>
<td>India</td>
<td>( 42) India</td>
<td>( 93) Singapore</td>
</tr>
<tr>
<td>Afghanistan</td>
<td>( 43) Afghanistan</td>
<td>( 94) Burkina</td>
</tr>
<tr>
<td>Thailand</td>
<td>( 44) Thailand</td>
<td>( 95) Zambia</td>
</tr>
<tr>
<td>Jamaica</td>
<td>( 45) Jamaica</td>
<td>( 96) Ecuador</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>( 46) Saudi Arabia</td>
<td>( 97) Uzbekistan</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>( 47) Ethiopia</td>
<td>( 98) No Nationality</td>
</tr>
<tr>
<td>Columbia</td>
<td>( 48) Columbia</td>
<td>( 99) Puerto Rico</td>
</tr>
<tr>
<td>Ghana</td>
<td>( 49) Ghana</td>
<td>(100) Laos</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>( 50) Bangladesh</td>
<td>(101) Estonia</td>
</tr>
<tr>
<td>Venezuela</td>
<td>( 51) Venezuela</td>
<td>(102) Angola</td>
</tr>
<tr>
<td>Tunisia</td>
<td>( 52) Tunisia</td>
<td>(103) Latvia</td>
</tr>
<tr>
<td>Mauritius</td>
<td>( 53) Mauritius</td>
<td>(104) Malaysia</td>
</tr>
<tr>
<td>Nigeria</td>
<td>( 54) Nigeria</td>
<td>(105) Namibia</td>
</tr>
</tbody>
</table>

**Var format**: NATION$$ (I2)

**19$$ - year : $$=84..11

**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. 449-30-89789-690)
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)

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 $PSBILB) 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, Ypbbla 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 = 'bil108.sav' /keep persnr hhnr yhnr 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 wz26 to wz30 wz46 to wz48
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
do if (Xp1001 ge 1 and Xp1001 le 4) Ypsedu = Xp1001
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
do if (Yb38 = 2)
end if

5. Ueberschreiben mit neu erworbenen Abschlüssen

if (Yp8701 ge 1 and Yp8701 le 4) Ypsedu = Yp8701
do if (Yp8701 eq 5)
end if

6. Integration von anderen Abschlüssen (_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 (Ypsedu = 1 and Ypsedua > 4) Ypsedu = 1
+ if (Ypsedu = 2 and Ypsedua > 4) Ypsedu = 2
+ if (Ypsedu = 3 and Ypsedua > 4) Ypsedu = 4
+ if (Ypsedu = 4 and Ypsedua > 4) Ypsedu = 5
+ end if

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

+ 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

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

+ 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
```plaintext
.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 (pbedu > 0)            Ypbeduo = pbedu
+ 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 Abschluessen ~~~
* ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
.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  (Y8706 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  =  Ypsedu
```
+ compute Ypbbila  =  Ypdbu
+ compute Ypsbilo  =  Ypsedu
+ compute Ypbbilo  =  Ypdbu

* For more information, contact: Peter Krause (Tel. +49-30-89789-690)
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..11

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

Var Label : $PSBILA Secondary school degrees/diplomas abroad
Value Label : (1) secondary school, no degree/diploma attained
              (2) secondary school, degree/diploma attained
              (3) vocational school
Var format : $PSBILA (I1)
$ - year : 84..11

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**

**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..11

**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

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 : $PBBILO  (11)

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)
$PBBIL02$

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..11$

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**

**Var Label**: $PBBIL03 *No vocational degree*

**Value Label**: $PBBIL03 (1) no vocational degree  
(2) apprenticeship  
(3) university  

**Var format**: $PBBIL03 (I1)

-$ year: 84..11

**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)
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..11  
Comment: As a supplement to the variable $PBBIL01, this variable gives (and updates) the highest-level vocational degree attained abroad.  

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

Var Label : $BILZEIT Amount of education or training (in years)
Var format : $BILZEIT (I4)
$ - year : 84..11

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:

\[
\begin{align*}
\text{school} &= -1 \quad *** \text{preset of intermediate schooling variable} *** \\
\text{occupa} &= -1 \quad *** \text{preset of intermediate occupational variable} *** \\
*** \text{computation for east-germany (mainly samples C and D)} *** \\
\text{ifthen} (\text{xPSBIL0} \geq 1 \text{ and school eq } -1) \\
\quad + \text{if } (\text{xPSBIL0} \text{ eq } 1) \text{ school } = 1 \\
\quad + \text{if } (\text{xPSBIL0} \text{ eq } 2) \text{ school } = 2 \\
\quad + \text{if } (\text{xPSBIL0} \text{ eq } 3) \text{ school } = 4 \\
\quad + \text{if } (\text{xPSBIL0} \text{ eq } 4) \text{ school } = 2 \\
\quad + \text{if } (\text{xPSBIL0} \text{ eq } 5) \text{ school } = 0 \\
\text{ifthen} (\text{xPBBILO} \geq 1 \text{ and occupa eq } -1) \\
\quad + \text{if } (\text{xPBBILO} \text{ eq } 1) \text{ occupa } = 1 \\
\quad + \text{if } (\text{xPBBILO} \text{ eq } 2) \text{ occupa } = 2 \\
\quad + \text{if } (\text{xPBBILO} \text{ eq } 3) \text{ occupa } = 3 \\
\quad + \text{if } (\text{xPBBILO} \text{ eq } 4) \text{ occupa } = 2 \\
*** \text{computation for foreigners (sample B and D)} **
*** \text{finished in their country of origin} **
\text{ifthen} (\text{xPSBILA} \geq 1 \text{ and school eq } -1) \\
\quad \text{if } (\text{xPSBILA} \text{ eq } 1) \text{ school } = 0 \\
\quad \text{if } (\text{xPSBILA} \text{ eq } 2) \text{ school } = 1 \\
\quad \text{if } (\text{xPSBILA} \text{ eq } 3) \text{ school } = 2 \\
\text{end if} \\
\text{ifthen} (\text{xPBBIILA} \geq 1 \text{ and occupa eq } -1) \\
\quad \text{.if } (\text{xPBBIILA} \text{ eq } 1) \text{ occupa } = 0 \\
\quad \text{.im } (\text{xPBBIILA} \text{ eq } 2) \text{ occupa } = 1 \\
\text{end if}

*** \text{All samples with finished education in Germany:} ****
\text{if } (\text{xPSBIL eq 6}) \text{ school } = 0 \\
\text{if } (\text{xPSBIL ge 1 and xPSBIL le 4}) \text{ school } = \text{xPSBIL} \\
\text{if } (\text{xPSBIL eq 5}) \text{ school } = 2 \\
\text{if } (\text{xspsbil eq 5 and xpsbilA eq 2}) \text{ school } = 1 \\
\text{if } (\text{xspsbil eq 5 and xpsbilA eq 3}) \text{ school } = 2 \\
\text{if } (\text{xPBBILO3 ge 1}) \text{ occupa } = 0 \\
\text{if } (\text{xPBBILO1 eq 1 or xPBBILO1 eq 5}) \text{ occupa } = 1 \\
\text{if } ((\text{xPBBILO1 ge 2 and xPBBILO1 le 4}) \text{ or xPBBILO1 eq 6}) \text{ occupa } = 2 \\
\text{if } (\text{xPBBILO2 eq 1}) \text{ occupa } = 3 \\
\text{if } (\text{xPBBILO2 ge 2}) \text{ occupa } = 4


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

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..11

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 ICED 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
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 |
    pbbil01==1 | pbbil01==2 | pbbil01==6 |
    pbilla==2 | pbilla==3 |
    pbilla==4 | pbilla==5 *3: middle voc.
isced=4 if(psbil==3 | psbil==4 | psbila==3) & (pbbil01==1 | pbbil01==2 | pbbil01==6 |
    pbilla==2 | pbilla==3) *4: voc. + Abi
isced=5 if pbbil01==3 | pbbil01==4 |
    (pbbil01==5 & (psbil==3 | psbil==4 | psbila==3)) *5: higher voc.
isced=6 if pbbil02==1 | pbbil02==2 | pbbil02==3 |
    pbbil02==4 | pbbil02==5 | pbbil02==6 *6: higher educ.
```
<table>
<thead>
<tr>
<th>ISCED</th>
<th>PSBIL Schulabschluss</th>
<th>PSBILA Schule im Ausland</th>
<th>PBBIL01 Berufliche Ausbildung</th>
<th>PBBILA Berufsbildung im Ausland</th>
<th>PBBIL02 Hochschulabschluss</th>
<th>ISCED97 – OECD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7. Noch kein Abschluss</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0. Noch kein Abschluss</td>
</tr>
<tr>
<td>1</td>
<td>5. Anderer Abschluss</td>
<td>1. Pflichtschule ohne Abschluss</td>
<td></td>
<td></td>
<td></td>
<td>1. Ohne Abschluss verlassen</td>
</tr>
</tbody>
</table>


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

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

**Value Label:**
- (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..11

**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 ($PSBIILO und $PBBILO) since only the West German version of the question has been asked since 1992.

The computation algorithm can be summarized as follows:

```plaintext
casmin=-1 if netold==1 | netold==5 | netold==7 *population
casmin= 0 if psbil==7 *0:0 in school
casmin= 1 if psbil==5 | 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) &
(psbila==2 | pbbila==3 |
(pbbila10>=1 & pbbila10<.)) *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 |
(pbbila10>=1 & pbbila10<.)) *5:2a intermediate vocational
casmin= 6 if psbil==3 | psbil==4) &
(pbbila==2 | pbbila==3 |
(pbbila10>=1 & pbbila10<.)) *6:2c_gen general maturity cer.
```

```plaintext
casmin= 7 if (psbil==3 | psbil==4) &
(pbbila==2 | pbbila==3 |
(pbbila10>=1 & pbbila10<.)) *7:2c_voc vocational maturity
```

```plaintext
casmin= 8 if pbbila02==1 *8:3a lower tertiary education
```

```plaintext
casmin= 9 if (pbbila10>=2 & pbbila10<5) | pbbila==4 *9:3b higher tertiary education
```

SOEP Survey Paper 111

SOEP v28
<table>
<thead>
<tr>
<th>CASMIN</th>
<th>PSBIL</th>
<th>Schulabschluss</th>
<th>PSBILA</th>
<th>Schule im Ausland</th>
<th>PBBIL01</th>
<th>Berufliche Ausbildung</th>
<th>PBBILA</th>
<th>Berufsbildung im Ausland</th>
<th>PBBIL02</th>
<th>Hochschulabschluss</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 0</td>
<td>7.</td>
<td>Noch kein Abschluss</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 – 1a</td>
<td>5.</td>
<td>Anderer Abschluss</td>
<td>6.</td>
<td>Ohne Abschluss verlassen</td>
<td>1.</td>
<td>Pflichtschule ohne Abschluss</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 – 1b</td>
<td>1.</td>
<td>Hauptschulabschluss</td>
<td>2.</td>
<td>Pflichtschule mit Abschluss</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 – 2a</td>
<td>1.</td>
<td>Realschulabschluss</td>
<td>2.</td>
<td>Weiterführende Schule</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 – 2c_gen</td>
<td>3.</td>
<td>Fachhochschulreife</td>
<td>4.</td>
<td>Abitur</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 – 3a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1. Fachhochschule</td>
</tr>
<tr>
<td>9 – 3b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Universitäten, TH</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. Hochschule im Ausland</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4. Ingenieur/Fachsch. Ost</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5. Hochschule (Ost)</td>
</tr>
</tbody>
</table>

Detailed description


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

Var Label : MONTH
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 : $84..11$

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 20\textsuperscript{th} of that month were recoded -3.

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

**Var Label** : MODE$$

**Interview method**

- (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 O onwards
- (200) with telephone assistance
- (210) written, by mail
- (220) telephone interview

**Var format** : MODE$$ (I2)

19$$ - year : $$=84..11

**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$$

Var Label : LABGRO$$  Current gross labor income in euros (generated)
Var format : LABGRO$$  (F2)
19$$ - year : $$=84..11

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$$**

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

**Var format** : IMPGRO$$ (I1)

19$$ - year : $$=84..11

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$$

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

Var format : LABNET$$ (F2)

19$$ - year : $$=84..11

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, wp5902, xp7302, yp6802, zp7202, bap6102

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

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

Var format : IMPNET$$ (I1)

19$$ - year : $$=84..11

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)
**FIELD$$**

Var Label : FIELD$$ “Field of tertiary education”
Value Label : 1-98 (see below)
Var Format : FIELD$$ (I2)

19$$-Year : $$=85..11

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sprach- und Kulturwissenschaften allgemein</td>
</tr>
<tr>
<td>2</td>
<td>Evang. Theologie, -Religionslehre</td>
</tr>
<tr>
<td>3</td>
<td>Kath. Theologie, -Religionslehre</td>
</tr>
<tr>
<td>4</td>
<td>Philosophie</td>
</tr>
<tr>
<td>5</td>
<td>Geschichte</td>
</tr>
<tr>
<td>6</td>
<td>Bibliothekswissenschaft, Dokumentation, Publizistik</td>
</tr>
<tr>
<td>7</td>
<td>Allgemeine und vergleichende Literatur und Sprachwissenschaft allgemein</td>
</tr>
<tr>
<td>8</td>
<td>Altphilologie (klass. Philologie), Neugriechisch</td>
</tr>
<tr>
<td>9</td>
<td>Germanistik (Deutsch, germanische Sprachen ohne Anglistik)</td>
</tr>
<tr>
<td>10</td>
<td>Anglistik, Amerikanistik</td>
</tr>
<tr>
<td>11</td>
<td>Romanistik</td>
</tr>
<tr>
<td>12</td>
<td>Slavistik, Baltistik, Finno-Ugristik</td>
</tr>
<tr>
<td>13</td>
<td>Aussereuropäische Sprach- und Kulturwissenschaften</td>
</tr>
<tr>
<td>14</td>
<td>Kulturwissenschaften i.e.S.</td>
</tr>
<tr>
<td>15</td>
<td>Psychologie</td>
</tr>
<tr>
<td>16</td>
<td>Erziehungswissenschaften</td>
</tr>
<tr>
<td>17</td>
<td>Sonderpädagogik</td>
</tr>
<tr>
<td>18</td>
<td>Sport, Sportwissenschaft</td>
</tr>
<tr>
<td>19</td>
<td>Rechts-, Wirtschafts- und Sozialwissenschaften allgemein</td>
</tr>
<tr>
<td>20</td>
<td>Regionalwissenschaften</td>
</tr>
<tr>
<td>21</td>
<td>Politikwissenschaften</td>
</tr>
<tr>
<td>22</td>
<td>Sozialwissenschaften</td>
</tr>
<tr>
<td>23</td>
<td>Sozialwesen</td>
</tr>
<tr>
<td>24</td>
<td>Rechtswissenschaften</td>
</tr>
<tr>
<td>25</td>
<td>Verwaltungswissenschaften</td>
</tr>
<tr>
<td>26</td>
<td>Wirtschaftswissenschaften</td>
</tr>
<tr>
<td>27</td>
<td>Wirtschaftsingenieurwesen</td>
</tr>
<tr>
<td>28</td>
<td>Mathematik, Naturwissenschaften allgemein</td>
</tr>
<tr>
<td>29</td>
<td>Mathematik</td>
</tr>
<tr>
<td>30</td>
<td>Informatik</td>
</tr>
<tr>
<td>31</td>
<td>Musik, Musikwissenschaft</td>
</tr>
<tr>
<td>32</td>
<td>Kunst, Kunstwissenschaft allgemein</td>
</tr>
<tr>
<td>33</td>
<td>Bildende Kunst</td>
</tr>
<tr>
<td>34</td>
<td>Gestaltung</td>
</tr>
<tr>
<td>35</td>
<td>Darstellende Kunst, Film und Fernsehen, Theaterwissenschaft</td>
</tr>
<tr>
<td>36</td>
<td>Musik, Musikwissenschaft allgemein</td>
</tr>
<tr>
<td>37</td>
<td>Ausserhalb der Studienbereichsgliederung</td>
</tr>
<tr>
<td>38</td>
<td>nicht zuzuordnen</td>
</tr>
</tbody>
</table>

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$$ contains more detailed information on the type of the degree. The data of the generated variable FIELD$$ 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 $P$.

**DEGREE$$**

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

Var Format : DEGREE$$ (I2)
19$$-Year : $$=85..11

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

In the variable DEGREE$$ we combine these two types of information. However, since the retrospective information was not collected before 2001 the variable covers until 2000 only persons for whom we have prospectively observed the end of study. This explains why the number of valid observations is rather small in these years. Information on the data source is stored in the variable FDT_F$$.
Each year the variable contains the most recently collected information. Take for instance a person for whom we have observed first an applied university diploma in 1987 and a university diploma in 1991. For this person the variables DEGREE$$ would be filled as follows:

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 $P.$

TRAINA$$

Var Label: TRAINA$$ “Apprenticeship – two-digit occupation KldB92”

Value Label : 1-99 (see below)

Var Format : TRAINA$$ (I2)
19$$-Year : $$=85..11

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 $P$.

**TRAINB$$**

Var Label: TRAINB$$ “Vocational school – two-digit occupation KldB92”

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

Var Format: TRAINB$$ (I2)

19$$-Year : $$=85..11

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$$**

Var Label : TRAINC$$ “Higher vocational school – two-digit occupation KldB92”

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

Var Format : TRAINC$$ (I2)

19$$-Year : $$=85..11

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$$**

Var Label : TRAIND$$ “Civil servant training – two-digit occupation KldB92”

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

Var Format : TRAIND$$ (I2)

19$$-Year : $$=85..11

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

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..10

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).
Variables no longer distributed:

ISCOUS$\$
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.