DIW Berlin
Deutsches Institut
für Wirtschaftsforschung

Data Documentation 7

2005

Markus M. Grabka

# Codebook for the $PEQUIV File 1984 – 2004

## CNEF variables with extended income information for the SOEP

DIW Berlin

## Data Documentation 7

Markus M. Grabka

# Codebook for the $PEQUIV File 1984 – 2004

CNEF variables with extended income information for the SOEP

Berlin, July 2005

* DIW Berlin, Längsschnittstudie "Sozio-oekonomisches Panel" (SOEP)
Homepage: http://www.diw.de/soep/
Contact: mgrabka@diw.de; phone: +49-30-897-89-339

IMPRESSUM



DIW Berlin
Deutsches Institut für Wirtschaftsforschung
Königin-Luise-Str. 5
14195 Berlin
Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
www.diw.de

ISSN 1861-1532

## PREFACE

The $PEQUV-File is based on the Cross-National Equivalent File (CNEF) with extended income information for the SOEP. This file comprises not only the aggregated income figures provided in the CNEF but also further single income components.

The CNEF is a joint effort of researchers and staff affiliated with Cornell University, the DIW Berlin, the University of Essex, and Statistics Canada funded by the National Institute on Aging and by the DIW Berlin. For extensive documentation of the CNEF cf. http://www.human.cornell.edu/pam/SOEP/equivfil.cfm or: Burkhauser, R.V, Butrica, B.A., Daly, M.C. and Lillard, D.R. (2001): The Cross-National Equivalent File: A product of cross-national research. In: Becker, I., Notburga, O and Rolf, G. (Eds.) Soziale Sicherung in einer dynamischen Gesellschaft (Social Insurance in a Dynamic Society). Festschrift für Richard Hauser zum 65. Geburtstag (Papers in Honor of the 65th Birthday of Richard Hauser). Campus, Frankfurt/New York.

**General notes:**

- In contrast to the original CNEF-data which is based on the 95% scientific use file of SOEP, the $PEQUIV-files include the full 100%-sample.
- The 2005 release of the $PEQUIV-files has been updated to include the 2004 (wave U) SOEP data..
- Population for $PEQUIV is made up by all members of households who were successfully interviewed (i.e., persons with $NETTO-codes 1 to 5 in the file PPFAD and $HNETTO-code 1 in the file HPFAD).
- For longitudinal consistency, all $PEQUIV income variables are consistently expressed in EURO (1 Euro = 1,95583 DM) independent of the currency used in the underlying survey instruments.
- Income data is missing for Sample C in 1990 and 1991 (first 2 waves of East German sample).
- Information for the SOEP sub-sample G ("High Income") has been incorporated for the survey years 2002-2004. However, the standard weighting schemes (i.e., variables w11101$$ - w11111$$) for these waves are based on SOEP sub-samples A-F only in order to maintain a harmonized time-series.
- The Consumer Price Index (Y11101$$) has been updated to the reference year 2000. Former differences in the Purchasing Power parity between East and West Germany are considered as well. Thus the variable Y11104$$ (Purchasing Power Parity for East Germany) has been dropped.

An important distinction from the original CNEF data, is that the $PEQUIV-files have been extended to also cover all single income components considered in the aggregated annual income figures of the CNEF. In principle, these single income components correspond to the originally surveyed information (which is stored in the $P, $PKAL and $H files, respectively) with some important amendments:

- Income variables are harmonized with respect to the periodicity, i.e. they give annual income (as of the previous calendar year). Components which are asked at monthly level have been multiplied by the number of months with receipt of the respective income (eventually, this implies imputation of missing number of months in the originally surveyed data as well as a longitudinally verified correction of implausible values).
- Any missing income information due to item-non response has been imputed according to the longitudinal and cross-sectional imputation procedures described in: Grabka, M.M. and Frick, J.R. (2003): Imputation of Item-Non-Response on Income Questions in the SOEP 1984–2002. DIW Research Note No. 29, Berlin October 2003. (http://www.diw.de/deutsch/produkte/publikationen/materialien/docs/papers/diw_rn03-10-29.pdf). See also: Frick, J.R. and Grabka, M.M. (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.
- An imputation flag for each of these single income components has been specified. These flags take a value of 1 if item-non-response on the underlying income variable has been imputed and 0 otherwise.

Variable naming conventions for the single income components: (see variable list on page 4):

- Variable names are longitudinally consistent using a two-digit suffix indicating the survey year (wave A = 84, wave B = 85, ..., wave U = 04).
- Variable names related to income components at the individual level start with the prefix “I”, e.g., Christmas bonus is given in variable IXMAS$$.
- The prefix “F” indicates the imputation flag, e.g. the flag variable for rental income (RENTY$$) is given by FRENTY$$.

For further information please contact: Markus M. Grabka (mgrabka@diw.de).

**Variables in the cross-sectional $PEQUIV Files 1984-2004**

## Using the $PEQUIV File Codebook

| | | |
|---|---|---|
| **Variable Name** | I1111084 - I1111004 | ***Name of Variable in the $PEQUIV-File*** |
| **Variable Label** | Labor Earnings of Individual | |
| **Unit of Observation** | Individual / Household / Year | |
| **Period** | Annual | ***Periodicity*** |
| | | ***Description of Variable Content*** |
| **Description** | This variable represents the labor earnings of each individual in the household. | |

**Method** ***Description of Variable Creation***

Labor earnings include wages and salary from all employment including training, primary and secondary jobs, and self-employment, plus income from bonuses, overtime, and profit-sharing.

Specifically labor earnings is the sum of income from primary job, secondary job, self-employment, 13th month pay, 14th month pay, Christmas bonus pay, holiday bonus pay, miscellaneous bonus pay, and profit-sharing income.

| | | |
|---|---|---|
| **Format** | Not formatted.<br>This variable is in current year EURO. | ***Variable Format in the $PEQUIV-File*** |

**Algorithm**

**1984-2004** algorithm: I11110$$ = IJOB1$$ + IJOB2$$ + ISELF$$ + IMILT$$ + I13LY$$ + I14LY$$ + IXMAS$$ + IHOLY$$ + IGRAY$$ + IOTHY$$ ($$=84-04)

| | |
|---|---|
| **Variable Name** | X11101LL |
| **Variable Label** | Unique Person Number |
| **Unit of Observation** | Individual |
| **Description** | This variable provides unique identification for each individual ever surveyed in the SOEP. |
| **Method** | Transcribed variable. |
| **Format** | N/A |
| | The original survey variables provided below can be found in the file PPFAD. |

**Algorithm**

**1984 – 2004** algorithm: X11101LL = PERSNR

| | |
|---|---|
| **Variable Name** | X1110284 - X1110204 |
| **Variable Label** | Household Identification Number |
| **Unit of Observation** | Household |
| **Description** | This variable links individuals to the households they were living in at the time of the interview. The SOEP provides yearly household identification numbers. |
| **Method** | Transcribed variable. |
| **Format:** | N/A |
| | The original survey variables provided below can be found in the files HPFAD. |
| **Algorithm** | |
| **1984-2004** algorithm: | X11102$$ = Yhhnr ($$=84-04, Y=A-U) |

**Variable Name** X1110384 - X1110304

**Variable Label** Individual in Household at Survey

**Unit of Observation** Individual

**Description** Indicates whether an individual was living in the household at the time of the interview regardless of whether the individual responded to interview questions.

**Method** Individuals who are members of a surveyed household are given a 1. Individuals who moved out of a SOEP household, died, or went abroad are given a 0.

**Format:**
0 = No
1 = Yes

The original survey variables provided below can be found in the file HPFAD.

**Algorithm**

**1984-2004** algorithm: If Ynetto = 1, 2, 3, 4, 5 and Yhnetto=1 then X11103$$ = 1 (Y=A-U, $$=84-04)
else X11103$$ = 0

**Variable Name** X11104LL

**Variable Label** Sub-sample Identifier

**Unit of Observation** Individual

**Description** This variable indicates from which sub sample an individual in the SOEP drawn.

**Method** The SOEP contains seven different samples.

1) Sample A consists of the residents of West Germany originally surveyed in 1984.
2) Sample B consists of a sample of foreign guest-workers of Italian, Spanish, Turkish, Yugoslavian, and Greek nationality also surveyed in 1984. The 1,400 foreign households in this original oversample were selected based on the nationality of the head of household. Foreigners who are not Italian, Spanish, Turkish, Yugoslavian, or Greek are included with the Germans in Sample A.
3) Sample C represents population of Germans residing in the former East Germans states in 1990.
4) Sample D surveyed in 1994/95 represents the population of households with at least household member who immigrated to Germany after 1984.
5) Sample E is a supplemental random sample that represents the population of Germany in 1998.
6) Sample F is an additional supplemental random sample that represents the population of Germany in 2000.
7) Sample G surveyed in 2002 is the so-called "high-income sample". Sampling criteria was a monthly disposable household income of more than 7.500 DM in 2001. Sample G represents about 7,3% individuals in private households with the highest income.

**Format**

21 = Sample A
22 = Sample B
23 = Sample C
24 = Sample D
25 = Sample E
26 = Sample F
27 = Sample G

The original survey variables provided below can be found in the file PPFAD.

**Algorithm**

**1984-2004** algorithm:

```
if psample = 1 then X11104LL = 21
else if psample = 2 then X11104LL = 22
else if psample = 3 then X11104LL = 23
else if psample = 4 then X11104LL = 24
else if psample = 5 then X11104LL = 25
else if psample = 6 then X11104LL = 26
else if psample = 7 then X11104LL = 27
```

**Variable Name** X1110584 - X1110503

**Variable Label** Indicator of Whether Person in Household was Interviewed

**Unit of Observation** Individual

**Description** Indicates whether an individual present in the household provided interview responses. Children in the household are counted as interviewed persons.

**Method** Individuals in the household 16 years of age and older who are members of a surveyed household reject to give an interview are given a 0.

**Format**
0 = Didn't provide information
1 = Provided information

The original survey variables provided below can be found in the file PPFAD.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: if Ynetto in (1,2,5) then X11105$$ = 1; (Y=A-U, $$=84-04)
else X11105$$ = 0;

| | |
|---|---|
| **Variable Name** | D1110184 - D1110104 |
| **Variable Label** | Age of Individual |
| **Unit of Observation** | Individual |
| **Description** | Indicates the age of the individual in years. |
| **Method** | The SOEP records the birth date (GEBJAHR) of each individual. The current age of an individual is created by subtracting the year of birth from the current year. |
| **Format** | -1 = Item non-response<br>0 = Newborn up to first birthday<br><br>The value of this variable ranges from 0 to 105.<br><br>The original survey variables provided below can be found in the file PPFAD. This algorithm omits individuals with survey non-responses. |
| **Algorithm** | |
| **1984-1999** algorithm: | D11101$$ = 19$$ - GEBJAHR ($$=84-99) |
| **2000-2004** algorithm: | D11101$$ = 20$$ - GEBJAHR ($$=00-04) |

| | |
|---|---|
| **Variable Name** | D11102LL |
| **Variable Label** | Gender of Individual |
| **Unit of Observation** | Individual |
| **Description** | Indicates the gender of the individual. |
| **Method** | The SOEP records the gender (SEX) of each individual. This information is acquired once and is not obtained in subsequent years.<br><br>Gender is constant through time and therefore does not have a yearly suffix. This variable is missing for the few cases where information about gender was not reported and inferences about gender could not be made. |
| **Format** | -1 = Item non-response<br>1 = Male<br>2 = Female<br><br>The original survey variables provided below can be found in the file PPFAD. This algorithm omits individuals with survey non-responses. |
| **Algorithm** | |
| **1984-2004** | algorithm: if SEX ne -1 then D11102LL = SEX |

| | |
|---|---|
| **Variable Name** | D1110384 - D1110304 |
| **Variable Label** | Race of Household Head |
| **Unit of Observation** | Individual |
| **Description** | Indicates the race of the interviewed head of household. |
| **Method** | Race is not available in the SOEP.<br><br>However, to separate Germans from non-Germans use the variables about<br>o nationality (NATION$$) which can be found in the $PGEN-files or<br>o the information about whether a person was born in Germany (GERMBORN) or<br>o the country of origin (CORIGIN) whereas both can be found in the PPFAD-file. |
| **Format** | -1 = no information available |

**Variable Name** D1110484 - D1110404

**Variable Label** Marital Status of Individual

**Unit of Observation** Individual

**Description** This variable indicates the marital status in the current survey year of all individuals in the household 16 years of age and older.

**Method** The married category represents individuals who are legally married and individuals who are living with a partner. All other categories represent the legal status of individuals who are not living with a partner. Married non-German "guest workers" whose spouses remained in their native countries are given a code of 6 or 7 depending on their ages.

**Format**

-1 = N/A – Child / Item non-response
1 = Married, Living with a Partner
2 = Single, not Living with a Partner
3 = Widowed, not Living with a Partner
4 = Divorced, not Living with a Partner
5 = Separated (Legally Married), not Living with a Partner
6 = Not Living with a Partner (Individuals age 18 and older)
7 = Not Living with a Partner (Individuals under age 18)

The original survey variables provided below can be found in the file _PGEN. This algorithm omits individuals with survey non-responses.

Equivalent Data File Variable Definitions: D11101__ = Age of Individual

**Algorithm**

**1984-2004** algorithm:

```
if D11101$$ ge 16 then do                    (Y=A-U, $$=84-04)
if      Yfamstd = 1 then D11104$$ = 1
else if Yfamstd = 2 then D11104$$ = 5
else if Yfamstd = 3 then D11104$$ = 2
else if Yfamstd = 4 then D11104$$ = 4
else if Yfamstd = 5 then D11104$$ = 3
else if Yfamstd = 6 and D11101$$ ge 18 then D11104$$ = 6
else if Yfamstd = 6 and D11101$$ lt 18 then D11104$$ = 7
end
if D11101$$ lt 16 then D11104$$ = -1
```

**Variable Name** D1110584 - D1110504

**Variable Label** Relationship to Household Head

**Unit of Observation** Individual

**Description** This variable indicates the individual's relationship to the current survey year's head of household.

**Method** The relation to head variable is created by collapsing the SOEP relationship to head variable into 5 categories. The original variable has 12 categories. These categories include spouses, life-partners, children, foster children, siblings, parents, in-laws, grandchildren, other relatives, and unrelated persons.

**Format**

-1 = Item non-response
1 = Head
2 = Partner
3 = Child
4 = Relative
5 = Non-relative

The original survey variables provided below can be found in the file _PBRUTTO. This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm:

if Ystell = 0 then D11105$$ = 1 (Y=A-U, $$=84-04)
else if Ystell = 1 or 2 then D11105$$ = 2
else if Ystell = 3 or 4 then D11105$$ = 3
else if Ystell = 5, 6, 7, 8, 9 or 10 then D11105$$ = 4
else if Ystell =11 then D11105$$ = 5
else D11104$$ = -1

**Variable Name** D1110684 - D1110604

**Variable Label** Number of Persons in Household

**Unit of Observation** Household

**Description** Indicates the number of persons in the household at the time of the interview.

**Method** This information is obtained from the household head or another household member who knows about the household's composition.

**Format** -1 = Item non-response
The value of this variable ranges from 1 to 17.

The original survey variables provided below can be found in the file _HBRUTTO. This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: D11106$$ = YHHGR (Y=A-U, $$=84-04)

**Variable Name** D1110784 - D1110704

**Variable Label** Number of Children in Household

**Unit of Observation** Household

**Description** Indicates the number of persons in the household under age 18 at the time of the interview.

**Method** This variable is created by computing the number of individuals in the household under the age of 18.

**Format** -1 = Item non-response
The value of this variable ranges from 0 to 10.

The original survey variables provided below can be found in the file PPFAD. This algorithm omits individuals with survey non-responses.

Equivalent Data File Variable Definitions: D11101__ = Age of Individual

**Algorithm**

**1984-2004** algorithm: if age$$ ge 0 and age$$ le 17 then sumkids$$=1
D11107$$ = sum of (sumkids$$) in the household ($$=84-04)

**Variable Name** D1110884 - D1110804

**Variable Label** Education With Respect to High School

**Unit of Observation** Individual

**Description** This variable indicates the highest level of education (less than high school, completed high school, or more than high school) of all individuals in the household 16 years of age and older.

**Method** This variable is coded as follows:

| | |
|---|---|
| Less than = High School | Intermediate secondary school (Realschule)<br>Lower secondary school (Hauptschule)<br>Other<br>None |
| High School = | Upper secondary school degree giving access to university studies Abitur) Certificate of aptitude for specialized short-course higher education (Fachhochschulreife)<br>Apprenticeship (Lehre)<br>Specialized vocational school (Berufsfachschule) |
| More than = High School | School of health care (Schule des Gesundheitswesens)<br>Specialized college of higher education, post-secondary technical (Fachhochschule)<br>College<br>Technical university usually requiring practical training as part of the studies (Technische Universität)<br>Civil service training |

**Format**

-1 = N/A – Child / Item non-response
1 = Less than High School
2 = High School
3 = More than High School

The original survey variables provided below can be found in the file _PGEN. This algorithm omits individuals with survey non-responses.

Equivalent Data File Variable Definitions: D11101__ = Age of Individual

**Algorithm**

**1984-2004** algorithm:

```
if Ypsbil=.B then Ypsbil=0;                    (Y=A-U, $$=84-04)
if Ypbbil01=.B then Ypbbil01=0;
if Ypbbil02=.B then Ypbbil02=0;
if Ypsbil in (1,2,5,6) then D11108$$=1;
if Ypsbil in (3,4) then D11108$$=2;
if Ypbbila in (3) then D11108$$=2;
if Ypbbila in (4) then D11108$$=3;
if Ypbbilo in (1) then D11108$$=2;
if Ypbbilo in (2,3,4) then D11108$$=3;
if Ypbbil01 in (1,2,4) then D11108$$=2;
if Ypbbil01 in (3,5) then D11108$$=3;
if Ypbbil02 in (1,2,3) then D11108$$=3;
if D11108$$ lt 0 then D11108$$=.M;
```

**Variable Name** D1110984 - D1110904

**Variable Label** Number of Years of Education

**Unit of Observation** Individual

**Description** This variable indicates the number of years of education completed at the time survey for all individuals in the household 16 years of age and older.

**Method** Individuals with a school leaving degree are assigned a minimum of between 9 and 12 years of education. Individuals with a vocational degree are assigned an additional 2 to 3.5 years of education. Individuals who attended a technical college are assigned an additional 4 years of education. If an individual received a vocational college degree or attended a university outside of the FRG then the individual is assigned a total of 18 or 19 years of education. For more information about the construction of this variable see Couch, 1994.

**Format** -1 = N/A – Child / Item non-response
The value of this variable ranges from 7 to 18.

The original survey variables provided below can be found in the file _PGEN. This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: if Ybilzeit=.B then D11109$$=0; (Y=A-U, $$=84-04)
else D11109$$ = Ybilzeit;
else D11109$$=-1;

| | |
|---|---|
| **Variable Name** | D11112LL |
| **Variable Label** | Race of individual |
| **Unit of Observation** | Individual |
| **Description** | Indicates the race of each individual in the household. |
| **Method** | Race is not available in the SOEP. |

**Variable Name** E1110180 - E1110104

**Variable Label** Annual Work Hours of Individual

**Unit of Observation** Individual

**Description** This variable reports annual hours worked in the previous year for all individuals in the household 16 years of age and older.

**Method** There is no direct report of annual work hours in the SOEP. This variable was constructed using information on employment status in the survey year, average number of hours worked per week, and the number of months worked in the previous year (reported in the activity calendar).

Annual hours worked in the previous year is calculated by adding together the estimated annual hours of full-time, part-time (including marginal employed), vocational training and short-time work. Annual hours of work in each of these four states is calculated by multiplying the average number of hours worked per week by the number of months worked in each of these three states for the previous year and by 4.33 (the average number of weeks per month). No correction for vacation or diseases has been made.

When the state indicated in the employment status variable matches the state recorded in the monthly calendar file we use the reported average number of hours worked per week as our measure of weekly hours worked.

When the state indicated in the employment status variable does <u>not</u> match the state recorded in the monthly calendar file we use an imputed weekly hours value as our measure of weekly hours worked. The imputed values are based on a regression of reported log average weekly work hours separately for full-time, part-time, marginal or irregular, short-time and currently not employed individuals. Covariates are age, age squared, sex, dummy variable for children in the household, marital status, region (east or west Germany) and two education dummy variables.

There are still some cases where no annual work hours but individual labor income (I11110$$) can be observed. This finding can be explained by e.g. individuals who took maternity leave or did National Service.

**Format** -1 = N/A - Child
0 = Not employed in the previous year
The value of this variable ranges from 1 to 7457.

The original survey variables provided below can be found in the file __P. This algorithm omits individuals with survey non-responses.

Equivalent Data File Variable Definitions: D11101__ = Age of Individual

**Algorithm**

**1984-2004** algorithm: if D11101$$ ge 16 then E11101$$=annual work hours imputation ($$=84-04)
else E11101$$=0

| | |
|---|---|
| **Variable Name** | E1120180 - E1120104 |
| **Variable Label** | Annual Work Hours of Individual Imputed |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates whether annual hours were imputed. |
| **Method** | This information is not available in the SOEP. |

**Variable Name** E1110284 - E1110204

**Variable Label** Employment Status of Individual

**Unit of Observation** Individual

**Description** This variable indicates employment status in the previous year for all individuals in the household 16 years of age and older.

**Method** If the individual had positive wages in the previous year and worked at least 52 hours then the individual was employed . Otherwise, the individual was not employed.

**Format**

-1 = N/A - Child
0 = Not Employed
1 = Employed

This algorithm omits individuals with survey non-responses.

Equivalent Data File Variable Definitions:
D11101__ = Age of Individual
E11101__ = Annual Work Hours of Individual
I11110__ = Individual Labor Earnings

**Algorithm**

**1984-2004** algorithm:

```
if D11101$$ ge 16 then do            ($$=84-04)
if I11110$$ gt 0 and E11101$$ ge 52 then E11102$$=1
else E11102$$=0
end
if D11101$$ lt 16 then E11102$$=-1
```

**Variable Name** E1110384 - E1110304

**Variable Label** Employment Level of Individual

**Unit of Observation** Individual

**Description** This variable indicates whether any individual in the household 16 years of age and older worked full-time, part-time, or not at all in the previous year.

**Method** If the individual had positive wages and worked at least 1,820 hours last year (35 hours per week on average), then the individual was employed full-time. If the individual had positive wages and worked at least 52 hours but less than 1,820 hours last year, then the individual was employed part-time. Otherwise, the individual was not working

**Format**

-1 = N/A - Child
1 = Full Time
2 = Part Time
3 = Not Working

This algorithm omits individuals with survey non-responses.

Equivalent Data File Variable Definitions:
D11101__ = Age of Individual
E11101__ = Annual Work Hours of Individual
E11102__ = Employment Status of Individual

**Algorithm**

**1984-2004** algorithm:

```
if D11101$$ ge 16 then do                                   ($$=84-04)
if E11102$$=1 then do
if E11101$$ ge 1,820 then E11103$$=1
else if E11101$$ ge 52 and E11101$$ lt 1,820 then E11103$$=2
end
if E11102$$=0 then E11103$$=3
end
if D11101$$ lt 16 then E11103$$=.-1
```

**Variable Name** E1110484 - E1110404

**Variable Label** Primary Activity of Individual

**Unit of Observation** Individual

**Description** This variable indicates primary activity at the time of the survey for all individuals in the household 16 years of age and older.

**Method** This variable is based on the individual's self-reported employment status at the time of the interview.
If the individual reported being full-time, part-time, or marginally employed, having short-time work, performing military/civilian service, on maternity leave, or being engaged in in-company training then the individual is considered to be working now.
If the individual reported not being employed or being unemployed then the individual is considered to be not working now.
Unemployed is not a category in the recoded variable because in the original data individuals were able to choose unemployed as their employment status in 1984 through 1990 only.

**Format**
-1 = N/A – Child
-2 = Item-non response
1 = Working Now
2 = Not Working Now

The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses.
Equivalent Data File Variable Definitions: D11101__ = Age of Individual

**Algorithm**

**1984-1990** algorithm:
if D11101$$ ge 16 or psample=3 then do
if VAR=1, 2, 3, 4 then E11104$$=1
else if VAR=5, 6 or 7 then E11104$$=2
else E11104$$=-2
end
else E11104$$=-1

(VAR=ap08, bp16, cp16, dp12, ep12 , fp10, gp12, zp16, $$=84-90)

**1991-1995** algorithm:
if D11101$$ ge 16 then do
if VAR=1, 2, 3, 4, 5, or 6 then E11104$$=1
else if VAR=7, 8 or 9 then E11104$$=2
else E11104$$=-2
end
else E11104$$=-1

(VAR=hp15, ip15, jp15, kp25, lp21, $$=91-95)

**1996-2004** algorithm:
if D11101$$ ge 16 then do
if VAR=1, 2, 3, 4 then E11104$$=1
else if VAR=5, 6, 7, 8 then E11104$$=2
else E11104$$=-2
end
else E11104$$=-1

(VAR=mp15, np11, op09, pp10, qp10, rp12, sp15, tp34, up09 $$=96-04)

**Variable Name** E1110584 - E1110504

**Variable Label** Occupation of Individual

**Unit of Observation** Individual

**Description** This variable indicates occupation at the time of the survey for all individuals in the household 16 years of age and older.

**Method** This variable is based on the individual's self-reported occupation at the time of the interview given by ISCO-88 occupation code (IS88$$). Occupation is coded as not applicable for individuals who were not working at the time of the interview. To harmonize the occupation of individual across countries we recalibrate the ISCO-88 code into the former ISCO68-code.

**Format**

-1 = N/A – Child
-2 = Item Non-response

| | | | |
|---|---|---|---|
| 0 = Not Applicable | 30 = Office Manager | 54 = Domestic Help | 80 = Shoemaker |
| 1 = Chemist | 31 = Administrator | 55 = Janitor | 81 = Cabinet Maker |
| 2 = Architect/Engineer | 32 = Stenographer | 56 = Dry-Cleaner | 82 = Stone Cutter |
| 3 = Eng.Tech.Expert | 33 = Bookkeepr/Cash | 57 = Hair Stylist | 83 = Tool/Die Maker |
| 4 = Aero/MarineEngr | 34 = ComputerOperat | 58 = SecurityServic | 84 = Machine Fitter |
| 5 = Life/Phys. Sci. | 35 = Transp. Attend | 59 = Service Worker | 85 = Electr. Enginr |
| 6 = Dr./Dentist/Vet | 36 = Conductor | 60 = Agriculturladm | 86 = Broadcaster |
| 7 = RelatMedicalJob | 37 = Mailman | 61 = Farm Manager | 87 = Pipe Fitter |
| 8 = Mathematician | 38 = Tel. Operator | 62 = Farm Hand | 88 = Jewelry Maker |
| 9 = Economist | 39 = Ofc.Worker Etc | 63 = Forestry Work | 89 = Glazier |
| 11 = Accountant | 40 = BusinessManagr | 64 = Fisher/Hunter | 90 = RubberProducer |
| 12 = Lawyer | 41 = Business Oper. | 70 = Inspector | 91 = Paper Producer |
| 13 = Educator | 42 = Buyer | 71 = Miner | 92 = Printer Etc. |
| 14 = Cleric | 43 = Tech.Salespers | 72 = Foundry Worker | 93 = Painter |
| 15 = Author | 44 = Insurance Rep. | 73 = Lumberman | 94 = Manufacturer |
| 16 = Sculptr/Paintr | 45 = Vendor | 74 = ChemicalWorker | 95 = Bricklay/Carpt |
| 17 = Music/Perform | 49 = Salesperson | 75 = Spinner/Weaver | 96 = Stat.Mach.Oper |
| 18 = Prof. Athlete | 50 = Rest./StoreMgr | 76 = Tanner/Furrier | 97 = Convey. Oper. |
| 19 = Scientist | 51 = Lodging Operat | 77 = Food Producer | 98 = Transport.Oper |
| 20 = Legislator | 52 = HH Supervisor | 78 = TobaccoProducr | 99 = Labor/Craftsmn |
| 21 = Priv.Bus.Leadr | 53 = Cook/Waiter | 79 = Tailor | 101 = Soldier |

The original survey variables provided below can be found in the file _PGEN. This algorithm omits individuals with survey non-responses. Equivalent Data File Variable Definitions: E11104$$ = Primary Activity of Individual

**Algorithm**

**1984-2004** algorithm: if X11103$$ = 1 then do ($$=84-04)

```
 if       E11104$$ in (2)                    then E11105$$=0;
 else if E11104$$ in (1) and is88$$ le 0 then E11105$$=.-1;
 else if E11104$$ in (1,2,3,4) and is88$$ gt 0 then do;
   if      is88$$ in (3,4,5,6,7,8,9,9920,9930,9941,9942,9950,9960,9970,9980,7216,7510,7520,7900)then E11105$$=.-1;
  else if is88$$ in (2113,2146,3111,3116)                                                      then E11105$$=1;
  else if is88$$ in (2139,2140,2141,2142,3112,3118)                                            then E11105$$=2;
  else if is88$$ in (2143,2144,2145,2147,2148,2149,3110,3113,3114,3115,3117,3119)              then E11105$$=3;
  else if is88$$ in (3141,3142,3143,3144,3145)                                                 then E11105$$=4;
  else if is88$$ in (2111,2112,2114,2211,2212,2213,3100,3211,3212,3213)                        then E11105$$=5;
  else if is88$$ in (2221,2222,2223)                                                           then E11105$$=6;
  else if is88$$ in (2224,2229,2230)                                                           then E11105$$=7;
  else if is88$$ ge 2121 and is88$$ le 2139                                                    then E11105$$=8;
  else if is88$$ in (2441)                                                                     then E11105$$=9;
  else if is88$$ in (2411)                                                                     then E11105$$=11;
  else if is88$$ in (2420,2421,2422,2429)                                                      then E11105$$=12;
  else if (is88$$ ge 2300 and is88$$ le 2359) or is88$$ in (3300,3310,3320,3330,3340)          then E11105$$=13;
  else if is88$$ in (2460,3480)                                                                then E11105$$=14;
  else if is88$$ in (2450,2451)                                                                then E11105$$=15;
```

```
  else if is88$$ in (2452,3131,3471,5151,5152)                                    then E11105$$=16;
  else if is88$$ in (2453,2454,2455,3470,3472,3473,3474,5210)                     then E11105$$=17;
  else if is88$$ in (3475)                                                        then E11105$$=18;
  else if is88$$ in (2000,2400,2412,2419,2431,2432,2442,2410,2443,2444,2445,2446,2500,3460)then E11105$$=19;
  else if is88$$ ge 1110 and is88$$ le 1143                                       then E11105$$=20;
  else if is88$$ in (1210,1220-1223,1226-1237,1239,1200,1300,1310,1312,1313,1316-1319)  then E11105$$=21;
  else if (is88$$ ge 3439 and is88$$ le 3449) or is88$$ in (2470,3450,3430,3432)  then E11105$$=31;
  else if is88$$ in (4111,4112,4113,4114,4115,4143,4144)                          then E11105$$=32;
  else if is88$$ in (3433,3434,4212,4121,4122,4211)                               then E11105$$=33;
  else if is88$$ in (5112)                                                        then E11105$$=36;
  else if is88$$ in (4142)                                                        then E11105$$=37;
  else if is88$$ in (4223)                                                        then E11105$$=38;
  else if is88$$ in (3431,4000,4190,3120,3121,4200,4400,4500)                     then E11105$$=39;
  else if is88$$ in (1224,1314)                                                   then E11105$$=41;
  else if is88$$ in (3416,3410,3422)                                              then E11105$$=42;
  else if is88$$ in (3415)                                                        then E11105$$=43;
  else if is88$$ in (3412,3411,3413,3421)                                         then E11105$$=44;
  else if is88$$ in (5220,5230,3414,3417,3419,3423,3429)                          then E11105$$=45;
  else if is88$$ in (3000)                                                        then E11105$$=49;
  else if is88$$ in (1225,1315)                                                   then E11105$$=51;
  else if is88$$ in (5120,5121,9141)                                              then E11105$$=52;
  else if is88$$ in (5122,5123)                                                   then E11105$$=53;
  else if is88$$ in (5130,5142,9131,9130)                                         then E11105$$=54;
  else if is88$$ in (7143,9132,9142)                                              then E11105$$=55;
  else if is88$$ in (9133)                                                        then E11105$$=56;
  else if is88$$ in (5141)                                                        then E11105$$=57;
  else if is88$$ in (3151,3152,5160,5161,5162,5163,5169,9152)                     then E11105$$=58;
  else if is88$$ in (5131-5133,5139,5143,5149,5111,5113,4213-4215,4221,4222,4131-4133,4141,
                     3221-3229,3231,3232, 3241,3242,4220,4300,5100)               then E11105$$=59;
  else if is88$$ in (1311)                                                        then E11105$$=60;
  else if is88$$ in (6111,6112,6113,6114,6121,6122,6123,6124,6129,6130,6120)      then E11105$$=61;
  else if is88$$ in (6100,9211,6210,9210)                                         then E11105$$=62;
  else if is88$$ in (6141,9212,9213,6142)                                         then E11105$$=63;
  else if is88$$ in (6150,6151,6152,6153,6154)                                    then E11105$$=64;
  else if is88$$ in (7111,7112,7113,8151,8111,8112,8113,7100,8110)                then E11105$$=71;
  else if is88$$ in (8120,8121,8122,8123,8124,8130,8131,8139)                     then E11105$$=72;
  else if is88$$ in (7421,8140,8141,8142,8143,8253,8286)                          then E11105$$=73;
  else if is88$$ in (8150,8159,8154,8155,8221,8229,8232,7344,8220)                then E11105$$=74;
  else if is88$$ in (8260,8264)                                                   then E11105$$=75;
  else if is88$$ in (7434,7437,7441,8265)                                         then E11105$$=76;
  else if is88$$ in (7410-7415,8270-8278)                                         then E11105$$=77;
  else if is88$$ in (7416,8279)                                                   then E11105$$=78;
  else if is88$$ in (7430,7431,7432,7433,7435,7436,8263,8261,8262,8269)           then E11105$$=79;
  else if is88$$ in (7442,7424,8266)                                              then E11105$$=80;
  else if is88$$ in (7420,7422,7423,8240)                                         then E11105$$=81;
  else if is88$$ in (7221,7222,7223,7224,7231,7232,7233,7311,7220)                then E11105$$=83;
  else if is88$$ in (7200,7210,7211,7212,7213,7214,7215,8211,8223,7341,7230)      then E11105$$=84;
  else if is88$$ in (7240,7241,7242,7243,7244,7245,7137,8282,8283)                then E11105$$=85;
  else if is88$$ in (3132)                                                        then E11105$$=86;
  else if is88$$ in (7136)                                                        then E11105$$=87;
  else if is88$$ in (7313,7312,7331,7332,7330)                                    then E11105$$=88;
  else if is88$$ in (7320,7321,7322,7323,7324,7342,7343)                          then E11105$$=89;
  else if is88$$ in (8230,8231)                                                   then E11105$$=90;
  else if is88$$ in (7345,7346,8250,8251,8252)                                    then E11105$$=92;
  else if is88$$ in (7141,7142)                                                   then E11105$$=93;
  else if is88$$ in (7120,7121,7122,7123,7124,7129,7131,7132,7133,7134,7135,7139) then E11105$$=95;
  else if is88$$ in (8152,8153,8160-8163,8171,8172,8212,8222,8224,8281,8280,8290,
                     8284,8285,3122,3123,3133,3139,8100,8200,8210)                then E11105$$=96;
  else if is88$$ in (8330)                                                        then E11105$$=97;
  else if is88$$ in (8311,8312,8320,8321,8322,8323,8324,8331,8332,8333)           then E11105$$=98;
  else if is88$$ in (7000,8000,8334,8340,9000,9110-9113,9120,9151,9153,9160-9162,9310-9313,
                     9320-9322,9330-9333,9910)                                    then E11105$$=99;
  else if is88$$ in (1,2,110)                                                     then E11105$$=101;
  else                                                                                 E11105$$=-1;
end;
else E11105$$=-2; end;
```

**Variable Name** E1110684 - E1110604

**Variable Label** 1 Digit Industry of Individual

**Unit of Observation** Individual

**Description** This variable indicates industry in which each individual in the household 16 years of age and older is employed at the time of the survey.

**Method** This variable is based on the individual's self-reported industry of occupation at the time of the interview. This variable is created by collapsing the SOEP industry variable into 10 broad categories. Industry is coded as not applicable for individuals who were not working at the time of the interview.

**Format**

-1 = N/A – Child / Item Non-response
0 = Not Applicable
1 = Agriculture
2 = Energy
3 = Mining
4 = Manufacturing
5 = Construction
6 = Trade
7 = Transport
8 = Bank/Insurance
9 = Services

The original survey variables provided below (NACE$$) can be found in the file _PGEN. This algorithm omits individuals with survey non-responses.

Equivalent Data File Variable Definitions: E11104$$ = Primary Activity of Individual

**Algorithm**

**1984-2004** algorithm:

```
if X11103$$ = 1 then do                                          ($$=84-04)
  if       E11104$$ in (5,6,7,8)                    then E11106$$=0;
  else if E11104$$ in (1,2,3,4) and nace$$ le 0     then E11106$$=-1;
  else if E11104$$ in (1,2,3,4) and nace$$ gt 0  then do;
            if       nace$$ in (1,2,5)              then E11106$$=1;
            else if nace$$ in (40,41)               then E11106$$=2;
            else if nace$$ in (10,11,12,13,14)      then E11106$$=3;
            else if nace$$ ge 15 and nace$$ le 37    then E11106$$=4;
            else if nace$$ in (96,97,100)           then E11106$$=4;
            else if nace$$ in (45)                  then E11106$$=5;
            else if nace$$ in (50,51,52)            then E11106$$=6;
            else if nace$$ in (60,61,62,63)         then E11106$$=7;
            else if nace$$ in (65,66,67)            then E11106$$=8;
            else if nace$$ ge 70 and nace$$ le 95    then E11106$$=9;
            else if nace$$ in (55,64,98,99)         then E11106$$=9;
            else E11106$$=-1;
  end;
else E11106$$=-2;
end;
```

**Variable Name** E1110784 - E1110704

**Variable Label** 2 Digit Industry of Individual

**Unit of Observation** Individual

**Description** This variable indicates industry in which each individual in the household 16 years of age and older is employed at the time of the survey.

**Method** This variable is based on the individual's self-reported industry of occupation at the time of the interview. Industry is coded as not applicable for individuals who were not working at the time of the interview.

**Format** The original survey variables provided below (NACE$$) can be found in the file _PGEN. This algorithm omits individuals with survey non-responses.

Equivalent Data File Variable Definitions:
E11104 __ = Primary Activity of Individual

**Algorithm**
**1984-2004** algorithm:

```
if X11103$$ = 1 then do                                                    ($$=84-04)
if E11104$$ in (5,6,7,8) then E11107$$=0;
else if E11104$$ in (1-4) and NACE$$ le 0 then E11107$$=-1;
else if E11104$$ in (1-) and NACE$$ gt 0 then do;
 if      NACE$$ in (15,96,97,100) then E11107$$=-1;
 else if NACE$$ in (1,2)          then E11107$$=1;
 else if NACE$$ in (3)            then E11107$$=2;
 else if NACE$$ in (40,41)        then E11107$$=3;
 else if NACE$$ in (10-14)        then E11107$$=4;
 else if NACE$$ in (23,24)        then E11107$$=5;
 else if NACE$$ in (25)           then E11107$$=6;
 else if NACE$$ in (26)           then E11107$$=7;
 else if NACE$$ in (27,28,37)     then E11107$$=8;
 else if NACE$$ in (29,34,35)     then E11107$$=9;
 else if NACE$$ in (30,31,32,33) then E11107$$=10;
 else if NACE$$ in (20,21,22,36) then E11107$$=11;
 else if NACE$$ in (17,18,19)     then E11107$$=12;
 else if NACE$$ in (16)           then E11107$$=13;
 else if NACE$$ in (45)           then E11107$$=14;
 else if NACE$$ in (50,51)        then E11107$$=16;
 else if NACE$$ in (52)           then E11107$$=18;
 else if NACE$$ in (60-64)        then E11107$$=21;
 else if NACE$$ in (65,66,67)     then E11107$$=22;
 else if NACE$$ in (55)           then E11107$$=24;
 else if NACE$$ in (90)           then E11107$$=25;
 else if NACE$$ in (73,80,92)     then E11107$$=27;
 else if NACE$$ in (85)           then E11107$$=28;
 else if NACE$$ in (70,71         then E11107$$=29;
 else if NACE$$ in (72,74,93,98) then E11107$$=30;
 else if NACE$$ in (91)           then E11107$$=31;
 else if NACE$$ in (95)           then E11107$$=32;
 else if NACE$$ in (75,99)        then E11107$$=33;
 else                                  E11107$$=-1
 end;
else E11107$$=-2;
end
```

| | | |
|---|---|---|
| **Variable Name** | H1110184 - H1110104 | Number of Household members age 0-14 |
| | H1110284 - H1110204 | Number of Household members age 15-18 |
| | H1110384 - H1110304 | Number of Household members age 0-1 |
| | H1110484 - H1110404 | Number of Household members age 2-4 |
| | H1110584 - H1110504 | Number of Household members age 5-7 |
| | H1110684 - H1110604 | Number of Household members age 8-10 |
| | H1110784 - H1110704 | Number of Household members age 11-12 |
| | H1110884 - H1110804 | Number of Household members age 13-15 |
| | H1110984 - H1110904 | Number of Household members age 16-18 |
| | H1111084 - H1111004 | Number of Household members age 19+ or 16-18 years old and independent |

**Unit of Observation** Household

**Description** These variables indicate the number of household members in the given age category living in the household at the time of the interview.

H11109$$ ($$=84-04) includes 16-18 year old youth who has not completed his or her Abitur and unmarried and living with a parent or married and separated and living with a parent.
H11110$$ ($$=84-04) includes 16-18 year old youth who have completed Abitur or is in college.

**Method** These variables are the simple count of all individuals in the household whose age is in the listed category.

**Format** The value of this variable ranges from 0 to 20.

The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: ($$=84-04)

```
*  CREATE IN COLLEGE VARIABLE *;

*First collapse variables for waves m-o;
mp14 = 0; if mp1402 = 1 then mp14=1; if mp1403 = 1 then mp14=2;
np10 = 0; if np1002 = 1 then np10=1; if np1003 = 1 then np10=2;
op03 = 0; if op0302 = 1 then op03=1; if op0303 = 1 then op03=2;

*Next create college attendance indicator;
array col{*} ap0502 bp1502 cp1502 dp1102 ep1102 fp0902 gp1102 hp0602 ip1402
       jp1402 kp1902 lp1502 mp14   np10   op03   pp0902 qp0902 rp1102
       sp1402 tp3302 up0802;
array collg{*} collg84 - collg$$;

do i = 1 to dim(col);
 if col{i} in (1,2,3) then collg{i}=1;
 else collg{i}=0;
end;

* CREATE AGE GROUP VARIABLE *;
array ak101{*}    age84    - age$$;
array ak102{*}    marst84  - marst$$;
array ak103{*}    insch84  - insch$$;
array ak104{*}    collg84  - collg$$;
array ak105{*}    hrel84   - hrel$$;
array age14{*}    age1484  - age14$$;
array age15_18{*} age1884  - age18$$;
```

```
array chld018{*}  chld1884 - chld18$$;
array age0_1{*}   age184   - age1$$;
array age2_4{*}   age484   - age4$$;
array age5_7{*}   age784   - age7$$;
array age8_10{*}  age1084  - age10$$;
array age11_12{*} age1284  - age12$$;
array age13_15{*} age1584  - age15$$;
array age16_18{*} age1684  - age16$$;
array adults{*}   adult84  - adult$$;

do i = 1 to dim(ad101);
 if ad101{i}=1 then do;
   age14{i}    = 0;
   age15_18{i} = 0;
   chld018{i}  = 0;
   age0_1{i}   = 0;
   age2_4{i}   = 0;
   age5_7{i}   = 0;
   age8_10{i}  = 0;
   age11_12{i} = 0;
   age13_15{i} = 0;
   age16_18{i} = 0;
   adults{i}   = 0;
   if  0 <= ak101{i} < 15 then age14{i}    = 1;
   if 15 <= ak101{i} < 19 then age15_18{i} = 1;
   chld018{i} = sum(age14{i},age15_18{i});

   *** Code up indicators for McClements household age distribution ***;;
   if  0 <= ak101{i} <  2 then age0_1{i}   = 1;
   if  2 <= ak101{i} <  5 then age2_4{i}   = 1;
   if  5 <= ak101{i} <  8 then age5_7{i}   = 1;
   if  8 <= ak101{i} < 11 then age8_10{i}  = 1;
   if 11 <= ak101{i} < 13 then age11_12{i} = 1;
   if 13 <= ak101{i} < 16 then age13_15{i} = 1;
   if 19 <= ak101{i}      then adults{i}   = 1;

   if 16 <= ak101{i} < 19  then do;
    if (ak102{i} ne 1 and ak103{i}= 1 and ak104{i}= 0 )
    or (ak102{i} =  5 and ak103{i}= 1 and ak104{i}= 0 and ak105{i}= 3)
      then age16_18{i}=1;

    if ak102{i}=1 or ak103{i}=0 or ak104{i}=1 or ak105{i} in (1,2)
      then adults{i}=1;
   end;
 end;
end;

*All variables are then summed by household id number (X11102$$) ($$=84-04)
```

**Variable Names** H1111284 - H1111204

**Variable Label** Indicator – Wife / spouse is in Household

**Unit of Observation** Household

**Description** These variables indicate the presence of a "wife or spouse" in the household.

**Method** These variables are simple indicator variables that a person who is a "wife/spouse" is present in a given year.

**Format**
0 = Not present
1 = Present

The variables provided below can be found in the $PEQUIV files. This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm:

```
if X11103$$=1 then do;                                   ($$=84-04)
                    H11112$$=0;
 if D11105$$=2 then  H11112$$=1;
end;
```

The following algorithms allow users to take Equivalent file variables and construct equivalence weights commonly used in various countries. To obtain equivalent household income, divide the equivalence scale weight into the household income variable. Here we present three typical example of equivalence weights:

| | |
|---|---|
| **Equivalence scale** | **OECD Equivalence Weights**<br>(referred to below as W11114$$) ($$=84-04) |
| **Unit of Observation** | Household |
| **Description** | Scale used by Organization for Economic Cooperation and Development (1982) |
| **Method** | Sets a single adult to be 1.0, each additional adult to be 0.7, and each child to be 0.5. |
| **Algorithm** | W11114$$=(1.0+0.7*(D11106$$-H11101$$-1)+0.5*H11101$$); |

---

| | |
|---|---|
| **Equivalence scale** | **Modified OECD Equivalence Weights**<br>(referred to below as W11115$$) ($$=84-04) |
| **Unit of Observation** | Household |
| **Description** | Scale used by Organization for Economic Cooperation and Development (1982), see also Aldi, et al. (1994). |
| **Method** | Sets a single adult to be 1.0, each additional adult to be 0.5, and each child to be 0.3. |
| **Algorithm** | W11115$$=(1.0+0.5*(D11106$$-H11101$$-1)+0.3*H11101$$); |

---

| | |
|---|---|
| **Equivalence scale** | **Other Equivalence Weights**<br>(e.g. Square root of the Household size referred to below as W11116$$) ($$=84-04) |
| **Unit of Observation** | Household |
| **Description** | Household equivalence weight based upon a single international scale. |
| **Method** | The weight is based upon a scale developed in Buhmann et al. (1988). The scale is characterized by the following equation: $EI = D/S^e$<br><br>Where equivalent income (El) equals total disposable household income (D) divided by household size (S) raised to the power (e). The parameter (e) represents the elasticity of the scale rate with respect to household size. Recent international studies on income inequality and poverty sponsored by the OECD (e.g., Forster 1990; Atkinson et al. 1994), and the Statistical Office of the European Commission (Hagenaars et al. 1994) and the Ruggles (1990) study of the United States use this type of exponential equivalence scale. We adopt a value of a equal to 0.5, which is most commonly used in international comparisons. |
| **Algorithm** | W11116$$=D11106$$**0.5; |

**Variable Name** L1110184 - L1110104

**Variable Label** State of Residence

**Unit of Observation** Household

**Description** This variable indicates the German federal state in which the household was located at the time of the survey

**Method** N/A

**Format**

-1 = Item non-response
0 = Berlin
1 = Schleswig-Holstein
2 = Hamburg
3 = Lower Saxony
4 = Bremen
5 = North-Rhine-Westfalia
6 = Hessen
7 = Rheinland-Pfalz, Saarland
8 = Baden Wuertemberg
9 = Bavaria
11 = Berlin (East)
12 = Mecklenburg-Vorpommern
13 = Branbenburg
14 = Saxony-Anhalt
15 = Thueringen
16 = Saxony

These states can be collapsed into regions. From 1984 through 1989 three regions can be defined to include the following states:

North:
Berlin (0), Schleswig-Holstein (1), Hamburg (2), Lower-Saxony (3), Bremen (4)
South:
Hessen (6), Baden-Wuertemberg (8), Bavaria (9)
West
North-Rhine-Westfalia (5), Rheinland-Pfalz, Saarland (7)

From 1990 to present four regions can be defined to include the following states:
North: Schleswig-Holstein (1), Hamburg (2), Lower-Saxony (3), Bremen (4)
South: Hessen (6), Baden-Wuertemberg (8), Bavaria (9)
West: North-Rhine -Westfalia (5), Rheinland-Pfalz, Saarland (7)
East: Berlin (0), Berlin East (11), Mecklenberg-Vorpommern (12), Brandenburg (13) Saxony-Anhalt (14), Thueringen (15), Saxony (16)

This algorithm omits individuals with survey non-responses.
Original variables below can be found in _HBRUTTO files

**Algorithm**

**1984-2004** algorithm: L11101$$=Ybula ($$=84-04, Y=A-U)

| | |
|---|---|
| **Variable Name** | L1110284 - L1110204 |
| **Variable Label** | Region of Residence |
| **Unit of Observation** | Household |
| **Description** | This variable indicates whether the household was located in the former East or West Germany at the time of the survey |
| **Method** | N/A |
| **Format** | -1 = Item non-response<br>1 = West Germany<br>2 = East Germany<br><br>This algorithm omits individuals with survey non-responses.<br><br>Original variables below can be found in _HBRUTTO files |
| **Algorithm** | |
| **1984-1989** algorithm: | L11102$$=1 |
| **1990-2004** algorithm: | L11102$$=Ysampreg ($$=90-04, Y=G-U) |

**Variable Name** Y1110184 - Y1110104

**Variable Label** Consumer Price Index

**Unit of Observation** Year

**Description** This variable provides consumer price indices necessary to convert current Euro amounts into constant Euro amounts. The base income year is 2000 (survey year 2001).

**Method** This value of this variable is derived from the DESTATIS 2003, Preise. Verbraucherpreisindex und Index der Einzelhandelspreise. Lange Reihe ab 1948 bis 2002. Basisjahr 2000.

To convert 1985 household labor income into 1992 euro, for example, multiply 1985 household labor income by the ratio of the 1992 consumer price index to the 1985 consumer price index.

Example: I1110385 * (Y1110192/Y1110185)

**Format** N/A

| | | **Price Index for West German States** | **Price Index for East German States** |
|---|---|---|---|
| **1984** | algorithm: | Y1110184 = 72.2 | N/A |
| **1985** | algorithm: | Y1110185 = 74.0 | N/A |
| **1986** | algorithm: | Y1110186 = 75.5 | N/A |
| **1987** | algorithm: | Y1110187 = 75.4 | N/A |
| **1988** | algorithm: | Y1110188 = 75.5 | N/A |
| **1989** | algorithm: | Y1110189 = 76.5 | N/A |
| **1990** | algorithm: | Y1110190 = 78.7 | N/A |
| **1991** | algorithm: | Y1110191 = 80.7 | N/A |
| **1992** | algorithm: | Y1110192 = 83.7 | Y1110192 = 70.4 |
| **1993** | algorithm: | Y1110193 = 87.0 | Y1110193 = 79.9 |
| **1994** | algorithm: | Y1110194 = 90.1 | Y1110194 = 88.3 |
| **1995** | algorithm: | Y1110195 = 92.6 | Y1110195 = 91.5 |
| **1996** | algorithm: | Y1110196 = 94.1 | Y1110196 = 93.3 |
| **1997** | algorithm: | Y1110197 = 95.3 | Y1110197 = 95.1 |
| **1998** | algorithm: | Y1110198 = 97.1 | Y1110198 = 97.2 |
| **1999** | algorithm: | Y1110199 = 97.9 | Y1110199 = 98.2 |
| **2000** | algorithm: | Y1110100 = 98.6 | Y1110100 = 98.6 |

**Price Index for German States**

| | | |
|---|---|---|
| **2001** | algorithm: | Y1110101 = 100.0 |
| **2002** | algorithm: | Y1110102 = 102.0 |
| **2003** | algorithm: | Y1110103 = 103.4 |

| | |
|---|---|
| **Variable Name** | I1110184 - I1110104 |
| **Variable Label** | Household Pre-Government Income |
| **Unit of Observation** | Household |
| **Period** | Annual |
| **Description** | This variable represents the combined income before taxes and government transfers of all individuals in the household 16 years of age and older. |
| **Method** | This variable is the sum of total family income from labor earnings, asset flows, private retirement income and private transfers. Labor earnings include wages and salary from all employment including training, self-employ-ment income, and bonuses, overtime, and profit-sharing. Asset flows include income from interest, dividends, and rent. Private transfers include payments from individuals outside of the household including alimony and child support payments. |
| **Format** | The value of this variable ranges from 0 to 5.000.000.<br>This variable is in current year EURO.<br><br>This algorithm omits individuals with survey non-responses.<br><br>Equivalent Data File Variable Definitions:<br>I11103__ = Household Labor Earnings<br>I11104__ = Household Asset Income<br>I11106__ = Household Private Transfers<br>I11117__ = Household Private Retirement Income |
| **Algorithm** | |
| **1984-2004** algorithm: | I11101$$ = I11103$$ + I11104$$ + I11106$$+I11117$$ ($$=84-04) |

**Variable Name** I1110284 - I1110204

**Variable Label** Household Post-Government Income

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the combined income after taxes and government transfers all individuals in the household 16 years of age and older.

**Method** This variable is the sum of total family income from labor earnings, asset flows, private retirement income, private transfers, public transfers, and social security pensions minus total family taxes. Labor earnings include wages and salary from all employment including training, self-employment income, bonuses, overtime, and profit-sharing. Asset flows include income from interest, dividends, and rent. Private transfers include payments from individuals outside of the household including alimony and child support payments. Public transfers include housing allowan-ces, child benefits, subsistence assistance from the Social Welfare Authority, special circumstances benefits from the Social Welfare Authority, government student assistance, maternity benefits, unemployment benefits, unemployment assistance, and unemployment subsistence allowance. Social security pensions include payments from old age, disability, and widowhood pension schemes. The tax burdens provided here are based upon updated and modified tax calculation routines developed by Schwarze. The tax burden includes income taxes and payroll taxes (health, unemployment, retirement insurance and nursing home insurance taxes). These routines are described in Schwarze (1995).

Household post-government income has no negative values.

**Format** The value of this variable ranges from 0 to 5.000.000.
This variable is in current year EURO.

This algorithm omits individuals with survey non-responses.

Equivalent Data File Variable Definitions:

I11103__ = Household Labor Earnings
I11104__ = Household Asset Income
I11106__ = Household Private Transfers
I11107__ = Household Public Transfers
I11108__ = Household Social Security Pensions
I11109__ = Total Household Taxes
I11117__ = Household Private Retirement Income

**Algorithm**

**1984-2004** algorithm:

I11102$$ = I11103$$ + I11104$$ + I11106$$ + I11107$$ + I11108$$ + I11117$$ - I11109$$ ($$=84-04)

| | |
|---|---|
| **Variable Name** | I1110384 - I1110304 |
| **Variable Label** | Household Labor Income |
| **Unit of Observation** | Household |
| **Period** | Annual |
| **Description** | This variable represents the combined labor income of all individuals in the household 16 years of age and older. |
| **Method** | Labor earnings include wages and salary from all employment including training, primary and secondary jobs, and self-employment, plus income from bonuses, overtime, and profit-sharing.<br><br>Specifically labor earnings is the sum of income from primary job, secondary job, self-employment, service pay, 13th month pay, 14th month pay, Christmas bonus pay, holiday bonus pay, miscellaneous bonus pay, and profit-sharing income. |
| **Format** | The value of this variable ranges from 0 to 5.000.000.<br>This variable is in current year EURO.<br><br>This algorithm omits individuals with survey non-responses.<br><br>Equivalent Data File Variable Definitions: I11110__ = Individual Labor Earnings |
| **Algorithm** | |
| **1984-2004** algorithm: | I11103$$ = sum of I11110$$ over all individuals in the household ($$=84-04) |

**Variable Name** I1110484 - I1110404

**Variable Label** Household Asset Income

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the household asset income reported by the head of the household.

**Method** Asset flows include income from interest, dividends, and rent.

After 1984 respondents who could not estimate their interest and dividend income directly were asked to select a range from a set of categories. Their choices were:
under 500 DM
500 to 2,000 DM
2,000 to 5,000 DM
5,000 to 10,000 DM
10,000 DM and over

Starting in year 2001 (wave R) an additional item was offered:
10,000 to 20,000 DM
20,000 DM and over

Since year 2002 (wave S) all items are asked for Euro:
under 250 Euro
250 to 1,000 Euro
1,000 to 2,500 Euro
2,500 to 5,000 Euro
5,000 to 10,000 Euro
10,000 Euro and over

These respondents are assigned an interest and dividend amount based on uniformly distributed random numbers within their income range.

Rental income is the amount of income from rent minus any operation and maintenance costs.

In 1991 income from rent and operation and maintenance costs were not asked. If the respondent was interviewed in 1990, 1991, and 1992 and reported having rental income or operation and maintenance costs for 1990 and 1992, the average of the 1990 and 1992 values are assigned to 1991. If the respondent was interviewed in only two of the years, one of the years being 1991, and reported having rental income or operation and maintenance costs, then rental income or operation and maintenance costs for that year are assigned to 1991.

**Format** The value of this variable ranges from 0 to 5.000.000.
This variable is in current year EURO.

The survey variables provided below are part of the $PEQUIV-file.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: I11104$$ = (RENTY$$ - OPERY$$) + DIVDY$$ ($$=84-04)

| | |
|---|---|
| **Variable Name** | I1110584 - I1110504 |
| **Variable Label** | Household Imputed Rental Value |
| **Unit of Observation** | Household |
| **Period** | Annual |
| **Description** | This variable represents the imputed rental value of owner occupied housing. |
| **Method** | The Imputed Rent (IR) information calculated for the German SOEP data is based on the so called Opportunity Cost Approach. This approach at the micro level yields information equivalent to that given by the Market Value Approach used in National account statistics for determining IR. After generating a fictitious market rent for owner-occupiers, all owner related costs are deducted including operating and maintenance costs, interest payments on mortgages, as well as property taxes (see Yates 1994 / United Nations 1977).<br>In more detail, the implementation of the opportunity cost approach is used here in the following way. Along the lines of Oaxaca (1973), we estimate an OLS (semilog) regression model of gross rent in terms of square meters (not including heating) actually paid by main tenants in privately financed housing (without social housing and households with reduced rent). Independent variables include indicators describing the condition of the house, the year of construction, size of dwelling, length of occupancy, community size and disposable income. Applying these regression coefficients to the population of owner occupiers yields an estimate of the gross value at market prices (without costs for heating and warm water). Further deducting owner-specific costs for taxation, maintenance and operating costs as well as interest on mortgages yields a net value which can be interpreted as the appropriate income advantage of owner-occupied housing. Information on interest and mortgage payments for the previous year from homeowners in SOEP serves as the basis for determining the level of interest payments. We assume an annuity with constant payments based on 7% annual interest and a 1% principal over the course of an average period of 30 years. In addition, we assume that mortgage payments begin at the same time in which the household moves into its new home. Thus, in the beginning of the repayment period interest payments clearly exceed the mortgage repayment. As times goes by, the share of the mortgage paid off increases, leaving an increasing income advantage from IR. For example an average interest burden of 3.29 DM/$m^2$ per month is used for West Germany in 1988 and rises to 5.52 DM/$m^2$ per month in 1998. The average interest burden in East Germany was slightly lower, at 4.14 DM/$m^2$ per month in 1998.<br>In case of owner related costs exceeding the income advantage (especially at the beginning of the mortgage repayment period), IR is assigned a value of zero.<br>For further details see: Frick and Grabka (2001) and Frick and Grabka (2003): |
| **Format** | The value of this variable ranges from 0 to 5.000.000.<br>This variable is in current year EURO. |
| **Algorithm** | N/a |

| | |
|---|---|
| **Variable Name** | I1110684 - I1110604 |
| **Variable Label** | Household Private Transfers |
| **Unit of Observation** | Household |
| **Period** | Annual |
| **Description** | This variable represents the combined private transfers of all individuals in the household 16 years of age and older. |
| **Method** | Private transfers consists of income received from persons outside of the interviewed household. For waves A-Q<br>this income was not specifically identified. Starting in wave R an additional question identifies alimony separately (variable $p2o03 in SOEP file $PKAL: $ = R, S, … ).<br>The bulk of transfer is likely to consist of alimony and child support payments. |
| **Format** | The value of this variable ranges from 0 to 5.000.000.<br>This variable is in current year EURO.<br><br>The survey variables provided below are part of the $PEQUIV-file.<br>This algorithm omits individuals with survey non-responses. |
| **Algorithm** | |
| **1984-2004** algorithm: | I11106$$ = sum of (IALM$$ + IELSE$$) over all individuals in the household ($$=84-04) |

**Variable Name** I1110784 - I1110704

**Variable Label** Household Public Transfers

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the combined public transfers of all individuals in the household 16 years of age and older.

**Method** Public Transfers are the sum of individual public transfers -- student grants, maternity benefits, unemployment benefits, unemployment assistance, subsistence allowance and transition pay -- over all individuals in the household, plus household benefits -- housing allowances, child benefits, nursing care insurance, direct housing subsidy , subsistence assistance, and support for special circumstances.

In 1984 the amount of child benefits is not asked. Child benefits for this year were imputed using information on the number of children in the household and the number of months the benefits were received.

In 1992 through 1994 the amounts of subsistence assistance and special circumstances benefits are not asked. These values have been filled in with imputed values for total social welfare income. In 1995 through 2000 amounts of subsistence assistance and special circumstances benefits are imputed using an algorithm developed by Peter Krause (DIW) based on the benefits received in the present survey month.

Since 1996 nursing care insurance benefits are included in the sum. In 1996 German law established direct housing subsidy payments. Starting in the 2000 survey a separate question was asked about income from this source. Direct housing subsidy payments for respondents who bought homes between 1996 and 1999 were imputed using information about the year of construction, acquisition of ownership and number of children in the household.

**Format** The value of this variable ranges from 0 to 5.000.000.
This variable is in current year EURO.

The survey variables provided below are part of the $PEQUIV-file.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: I1110704 = [sum of (IUNBY$$ + IUNAY$$ + ISUBY$$ + IERET$$ + IMATY$$ + ISTUY$$) over all individuals in the household] + HOUSE$$ + CHSPT$$ + NURSH$$ + SUBST$$ + SPHLP$$ + HSUP$$ ($$=84-04)

**Variable Name** I1110884 - I1110804

**Variable Label** Household Social Security Pensions

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the combined social security pensions of all individuals in the household 16 years of age and older.

**Method** Social security pensions are the sum of old-age, disability, and widowhood social security pensions. This include payments of the German Pension Insurance (GRV), Miner's social Insurance (Knappschaft), Civil Servant Pension (Beamtenpension), War Victim Benefits (Kriegsopferversorgung), Farmer's Benefits and accident pension (GUV).

In 1993 through 1994 pension income from East German pensions ($p7902o and $p7912o) is assigned to other pension income.

**Format** The value of this variable ranges from 0 to 5.000.000.
This variable is in current year EURO.

The survey variables provided below are part of the $PEQUIV-file.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985, 2002** algorithm:

I11108$$ = sum of (IOLDY$$ + IWIDY$$ + ICOMP$$ + IPRVP$$) over all individuals in the household ($$=84-85, 02)

**1986-2001, 2004** algorithm:

I11108$$ = sum of (igrv1$$ + igrv2$$ + ismp1$$ + ismp2$$ + iciv1$$ + iciv2$$ + iwar1$$ + iwar2$$ + iagr1$$ + iagr2$$ + iguv1$$ + iguv2$$) over all individuals in the household ($$=86-04)

| | |
|---|---|
| **Variable Name** | I1110984 - I1110903 |
| **Variable Label** | Total Household Taxes |
| **Unit of Observation** | Household |
| **Period** | Annual |
| **Description** | This variable includes income taxes and payroll taxes (e.g. health, unemployment, nursing home and retirement insurance taxes) of all individuals in the household 16 years of age and older. |
| **Method** | The tax estimates come from Schwarze (1995), the taxes are assigned on a household basis. The estimated tax burdens include income taxes and payroll taxes (health, unemployment, care and retirement insurance taxes). These routines are described in Schwarze (1995). Since 1995 the solidarity surplus tax is also considered in the tax estimates. No algorithms are provided for the tax estimates. |
| **Format** | The value of this variable ranges from 0 to 5.000.000.<br>This variable is in current year EURO.<br><br>The survey variables provided below are part of the $PEQUIV-file. |
| **Algorithm** | |
| **1984-2004** algorithm: | I11109$$ = I11111$$ + I11112$$ ($$=84,85,…,04) |

**Variable Name** I1111084 - I1111004

**Variable Label** Individual Labor Earnings

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents the labor earnings of individuals in the household 16 years of age and older.

**Method** Labor earnings include wages and salary from all employment including training, primary and secondary jobs, and self-employment, plus income from bonuses, overtime, and profit-sharing.

Specifically labor earnings is the sum of income from primary job, secondary job, self-employment, 13th month pay, 14th month pay, Christmas bonus pay, holiday bonus pay, miscellaneous bonus pay, and profit-sharing income.

Since 1996 military service payments are also considered.

**Format** The value of this variable ranges from 0 to 5.000.000.
This variable is in current year EURO.

The survey variables provided below are part of the $PEQUIV-file.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: I11110$$ = IJOB1$$ + IJOB2$$ + ISELF$$ + IMILT$$ + I13LY$$ + I14LY$$ + IXMAS$$ + IHOLY$$ + IGRAY$$ + IOTHY$$ ($$=84-04)

| | |
|---|---|
| **Variable Name** | I1111184 - I1111104 |
| **Variable Label** | Household Federal Taxes |
| **Unit of Observation** | Household |
| **Period** | Annual |
| **Description** | This variable includes federal income taxes of all individuals in the household 16 years of age and older. |
| **Method** | The tax estimates come from Schwarze (1995). Taxes are estimated for each tax unit within the household and then summed over all tax units within the household to arrive at a total household tax burden. The estimated tax burdens include federal income taxes and solidarity surplus tax. These routines are described in Schwarze (1995). No algorithms are provided for the tax estimates. |
| **Format** | The value of this variable ranges from 0 to 5.000.000.<br>This variable is in current year EURO. |
| **Algorithm** | N/a |

| | |
|---|---|
| **Variable Name** | I1111284 - I1111204 |
| **Variable Label** | Household Social Security Taxes |
| **Unit of Observation** | Household |
| **Period** | Annual |
| **Description** | This variable includes social security taxes (payroll taxes) of all individuals in the household 16 years of age and older. |
| **Method** | The tax estimates come from Schwarze (1995). Taxes are estimated for each tax unit within the household and then summed over all tax units within the household to arrive at a total household tax burden. The estimated tax burdens include social security taxes (e.g. health, unemployment, nursing home and retirement insurance taxes). These routines are described in Schwarze (1995). No algorithms are provided for the tax estimates. |
| **Format** | The value of this variable ranges from 0 to 5.000.000.<br>This variable is in current year EURO. |
| **Algorithm** | N/a |

| | | |
|---|---|---|
| **Variable Name** | I1111384 - I1111304 | Household Post-Government Income (TAXSIM) |
| | I1111484 - I1111404 | Total Household Taxes (TAXSIM) |
| | I1111584 - I1111504 | Household State Taxes (TAXSIM) |
| | I1111684 - I1111604 | Household Federal Taxes (TAXSIM) |

**Unit of Observation** Household

**Description** This variable represents the combined income after taxes and government transfers, the Total Household Taxes, the Household State Taxes and the Household Federal Taxes of all individuals in the household 16 years of age and older.

**Method** Income taxes and state taxes were not estimated for the SOEP using the National Bureau of Economic Research (NBER) TAXSIM Model. This variable is not available in the SOEP.

**Format** N/A

**Algorithm** N/a

**Variable Name** I1111784 - I1111704

**Variable Label** Household Private Retirement Income

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the combined retirement income from private sources of all individuals in the household 16 years of age and older.

**Method** Private pension income is the sum of supplementary civil servant pension income, company pensions, private pensions and pension income from "other" sources. See the algorithm for I11108$$.

**Format** The value of this variable ranges from 0 to 5.000.000.
This variable is in current year EURO.

The survey variables provided below are part of the $PEQUIV-file.
This algorithm omits individuals with survey non-responses.
Note also that this information is not available in 1984 and 1985.

**Algorithm:**

**1984-1985**: algorithm: N/a

**2002-2003** algorithm: I11117$$ = sum of (ICOMP$$ + IPRVP$$) over all individuals in the household
($$= 02-03)

**1986-2001, 2004** algorithm:
I11117$$ = sum of (ivbl1$$ + ivbl2$$ + icom1$$ + icom2$$ + iprv1$$ + iprv2$$
+ ison1$$ + ison2$$) over all individuals in the household
($$= 86-01, 04)

**Variable Name** I1111884 - I1111804

**Variable Label** Household Windfall Income

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the amount of total household windfall income of all individuals in the household 16 years of age and older.

**Method** Windfall income consists of one-time transfers, winnings, inheritance and gifts of money or items worth more than 5000 DM (wave Q-R) or 2500 Euro (wave S-T) . It was asked the first time in wave Q (variable qh4505 in SOEP file QH).

**Format** The value of this variable ranges from 0 to 5.000.000.
This variable is in current year EURO.

This algorithm omits individuals with survey non-responses.
The original survey variables provided below can be found in the file _H.

**Algorithm**

**2000- 2004** algorithm: I11118$$ = Ywind ($$=00-04)

Windfall income variable list by survey year - each entry denoted in algorithm as Ywind: QH4505 RH4505 SH4505 TH44 UH44

**Variable Name** I1120184 - I1120104

**Variable Label** Impute Household Pre-Government Income

**Unit of Observation** Household

**Description** This variable indicates the percentage share of income that has been imputed for Household Pre-Government Income.
The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003).

**Method** In the original SOEP data there are three types of missing values. These missing values can be interpreted as:

-1 = no answer or do not know
-2 = does not apply
-3 = original value was deleted because it was found to be implausible

The imputation procedures was used to fill in missing values represented by -1 and -3 only.

The components of household pre-government income (I11101$$) are household labor income, household asset income, household private retirement income, and household private transfers. These components can be found in the Equivalent Data File.

**Format** 0 = Not Imputed
1 = Fully Imputed

This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: I11201$$= percentage share of income that has been imputed for I11101$$ ($$=84-04)

**Variable Name** I1120284 - I1120204

**Variable Label** Impute Household Post-Government Income

**Unit of Observation** Household

**Description** This variable indicates the percentage share of income that has been imputed for household post-government income.
The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003).

**Method** In the original SOEP data there are three types of missing values. These missing values can be interpreted as:

-1 = no answer or do not know
-2 = does not apply
-3 = original value was deleted because it was found to be implausible

The imputation procedures was used to fill in missing values represented by -1 and -3 only.

The components of household post-government income (I11102$$) are household labor income, household asset income, household private retirement income, household private transfers, household public transfers, household social security pensions, and household taxes. These components can be found in the Equivalent Data File.

**Format** 0 = Not Imputed
1 = Fully Imputed

This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: I11202$$= percentage share of income that has been imputed for I11102$$ ($$=84-04)

**Variable Name** I1120384 - I1120304

**Variable Label** Impute Household Labor Income

**Unit of Observation** Household

**Description** This variable indicates the percentage share of income that has been imputed for household labor income.
The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003).

**Method** In the original SOEP data there are three types of missing values. These missing values can be interpreted as:

-1 = no answer or do not know
-2 = does not apply
-3 = original value was deleted because it was found to be implausible

The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.

The components of household labor income (I11103$$) are income from primary job, income from secondary job, income from self-employment, service pay, 13th month pay, 14th month pay, Christmas bonus pay, holiday bonus pay, miscellaneous bonus pay, and profit-sharing income. These components can be found in the SOEP.

**Format** 0 = Not Imputed
1 = Fully Imputed

This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: I11203$$= percentage share of income that has been imputed for I11103$$ ($$=84-04)

**Variable Name** I1120484 - I1120404

**Variable Label** Impute Household Asset Income

**Unit of Observation** Household

**Description** This variable indicates the percentage share of income that has been imputed for household asset income.
The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003).

**Method** In the original SOEP data there are three types of missing values. These missing values can be interpreted as:

-1 = no answer or do not know
-2 = does not apply
-3 = original value was deleted because it was found to be implausible

The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.

The components of household asset income (I11104$$) are interest and dividend income and rental income. These components can be found in the SOEP.

**Format** 0 = Not Imputed
1 = Fully Imputed

This algorithm omits individuals with survey non-responses.

Equivalent Data File Variable Definitions: I11104__ = Household Asset Income

**Algorithm**

**1984-2004** algorithm: I11204$$= percentage share of income that has been imputed for I11104$$ ($$=84-04)

| | |
|---|---|
| **Variable Name** | I1120584 - I1120504 |
| **Variable Label** | Impute Household Imputed Rental Value |
| **Unit of Observation** | Household |
| **Description** | This variable indicates if the imputed rental value has been imputed. |
| **Method** | Household imputed rental value is fully simulated in the SOEP. Thus this variable indicates if household Imputed rental Value has been simulated and takes a value of 0 otherwise. For further details see: Frick and Grabka (2001) and Frick and Grabka (2003). |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses.<br><br>Equivalent Data File Variable Definitions:<br>I11105$$ = Household Imputed Rental Value |
| **Algorithm** | |
| **1984-2004** algorithm: | I11205$$= 1 if I11105$$ > 0, else I11205$$ = 0. ($$=84-04) |

**Variable Name** I1120684 - I1120604

**Variable Label** Impute Household Private Transfers

**Unit of Observation** Household

**Description** This variable indicates the percentage share of income that has been imputed for household private transfers.
The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003).

**Method** In the original SOEP data there are three types of missing values. These missing values can be interpreted as:

-1 = no answer or do not know
-2 = does not apply
-3 = original value was deleted because it was found to be implausible

The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.

The components of household private transfers (I11106$$) are income received from outside of the household and alimony. These components can be found in the SOEP.

**Format** 0 = Not Imputed
1 = Fully Imputed

This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: I11206$$= percentage share of income that has been imputed for I11106$$ ($$=84-04)

**Variable Name** I1120784 - I1120704

**Variable Label** Impute Household Public Transfers

**Unit of Observation** Household

**Description** This variable indicates the percentage share of income that has been imputed for household public transfers.
The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003).

**Method** In the original SOEP data there are three types of missing values. These missing values can be interpreted as:

-1 = no answer or do not know
-2 = does not apply
-3 = original value was deleted because it was found to be implausible

The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.

The components of household public transfers (I11107$$) are government student assistance, maternity benefits, unemployment benefits, unemployment assistance, unemployment subsistence allowance, housing allowances, child benefits, subsistence assistance from the Social Welfare Authority, special circumstances benefits from the Social Welfare Authority, nursing home support payments, and direct housing subsidy payments. These components can be found in the SOEP.

**Format** 0 = Not Imputed
1 = Fully Imputed

This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: I11207$$= percentage share of income that has been imputed for I11107$$ ($$=84-04)

**Variable Name** I1120884 - I1120804

**Variable Label** Impute Household Social Security Pensions

**Unit of Observation** Household

**Description** This variable indicates the percentage share of income that has been imputed for household social security pensions
The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003).

**Method** In the original SOEP data there are three types of missing values. These missing values can be interpreted as:

-1 = no answer or do not know
-2 = does not apply
-3 = original value was deleted because it was found to be implausible

The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.

The components of household social security pensions (I11108$$) are social security pensions for old-age, disability, and widowhood. These components can be found in the SOEP.

**Format** 0 = Not Imputed
1 = Fully Imputed

This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: I11208$$= percentage share of income that has been imputed for I11108$$ ($$=84-04)

| | |
|---|---|
| **Variable Name** | I1120984 - I1120904 |
| **Variable Label** | Impute Total Household Taxes |
| **Unit of Observation** | Household |
| **Description** | Total household taxes are fully simulated in the SOEP. Thus this variable indicates if total household taxes has been simulated and takes a value of 0 otherwise. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses.<br><br>Equivalent Data File Variable Definitions: I11109__ = Total Household Taxes |
| **Algorithm** | |
| **1984-2004** algorithm: | I11209$$= 1 if I11109$$ > 0, else I11209$$ = 0. ($$=84-04) |

**Variable Name** I1121084 - I1121004

**Variable Label** Impute Individual Labor Earnings

**Unit of Observation** Individual

**Description** This variable indicates the percentage share of income that has been imputed for individual labor earnings.
The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003).

**Method** In the original SOEP data there are three types of missing values. These missing values can be interpreted as:

-1 = no answer or do not know
-2 = does not apply
-3 = original value was deleted because it was found to be implausible

The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.

The components of individual labor income (I11110$$) are income from primary job, income from secondary job, income from self-employment, service pay, 13th month pay, 14th month pay, Christmas bonus pay, holiday bonus pay, miscellaneous bonus pay, and profit-sharing income. These components can be found in the SOEP.

**Format** 0 = Not Imputed
1 = Fully Imputed

This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: I11210$$= percentage share of income that has been imputed for I11110$$ ($$=84-04)

**Variable Name** I1121784 - I1121704

**Variable Label** Impute Household Private Retirement Income

**Unit of Observation** Household

**Description** This variable indicates the percentage share of income that has been imputed for household private retirement income.
The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003)

**Method** In the original SOEP data there are three types of missing values. These missing values can be interpreted as:

-1 or .A = no answer or do not know
-2 or .B = does not apply
-3 or .C = original value was deleted because it was found to be implausible

The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.

The components of private pension income (I11117$$) are supplementary civil servant pension income, company pensions, private pensions and pension income from "other" sources. These components can be found in the SOEP.

**Format** 0 = Not Imputed
1 = Fully Imputed

This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: I11217$$= percentage share of income that has been imputed for I11117$$ ($$=84-04)

**Variable Name** I1121884 - I1121804

**Variable Label** Impute Household Windfall Income

**Unit of Observation** Household

**Description** This variable indicates if windfall income (I11118$$) has been imputed.
Item non response on windfall income has been imputed using the sample median in the respective observation years. For further details, see: Grabka and Frick (2003)

**Method** In the original SOEP data there are three types of missing values. These missing values can be interpreted as:

-1 or .A = no answer or do not know
-2 or .B = does not apply
-3 or .C = original value was deleted because it was found to be implausible

The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.

**Format** 0 = Not Imputed
1 = Fully Imputed

This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: N/a

**Variable Name** RENTY84 – RENTY04

**Variable Label** Income from rental and leasing

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the household income from rental and leasing.

In 1991 income from rent were not asked. If the respondent was interviewed in 1990, 1991, and 1992 and reported having rental income for 1990 and 1992, the average of the 1990 and 1992 values are assigned to 1991. If the respondent was interviewed in only two of the years, one of the years being 1991, and reported having rental income, then rental income for that year are assigned to 1991.

**Method** Transcribed variable.

**Format** The value of this variable ranges from 0 to 999.999.

This variable is in current year EURO.

The original survey variables provided below can be found in the file _H.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm:

```
array cm101{*} ah41    bh35    ch47    dh47    eh38    fh38    gh38
               hhrenty ih41    jh41    kh41    lh41    mh41    nh41
               oh41    ph41    qh41    rh41    sh41    th39    uh39;
*** imputed values due to item-non response ***
array cm102{*} xah41   xbh35   xch47   xdh47   xeh38   xfh38   xgh38
               out     xih41   xjh41   xkh41   xlh41   xmh41   xnh41
               xoh41   xph41   xqh41   xrh41   xsh41   xth39   xuh39;
array cm103{*} temp184 - temp104;
array cm104{*} renty84 - renty04;

do i = 1 to dim(netto);
 cm103{i}=.;
 if netto{i}=1 then do;
  if cm102{i} lt 0 then cm102{i} = 0;
  if cm101{i} eq .A or cm101{i} eq .C then cm103{i}=cm102{i};
  else if cm101{i} eq .B then cm103{i}=0;
  else cm103{i}=cm101{i};
  cm104{i}=cm103{i};
 end;
 else cm104{i}=-2;
end;
```

**Variable Name** OPERY84 – OPERY04

**Variable Label** Operation, maintenance costs

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the household operation and maintenance costs

In 1991 operation and maintenance costs were not asked. If the respondent was interviewed in 1990, 1991, and 1992 and reported having operation and maintenance costs for 1990 and 1992, the average of the 1990 and 1992 values are assigned to 1991. If the respondent was interviewed in only two of the years, one of the years being 1991, and reported having operation and maintenance costs, then operation and maintenance costs for that year are assigned to 1991.

**Method** Transcribed variable.

**Format** The value of this variable ranges from 0 to 999.999.

This variable is in current year EURO.

The original survey variables provided below can be found in the file _H.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm:

```
array ct101{*} ah4201  bh3601  ch4801  dh4801  eh3901  fh3901  gh3901
               hhopery ih4201  jh4201  kh4201  lh4201  mh4201  nh4201
               oh4201  ph4201  qh4201  rh4201  sh4201  th4001  uh4001;
*** imputed values due to item-non response ***
array ct102{*} xah4201 xbh3601 xch4801 xdh4801 xeh3901 xfh3901 xgh3901
               out     xih4201 xjh4201 xkh4201 xlh4201 xmh4201 xnh4201
               xoh4201 xph4201 xqh4201 xrh4201 xsh4201 xth4001 xuh4001;
array ct103{*} temp184 - temp104
array ct104{*} opery84 - opery04;

do i = 1 to dim(netto);
 ct103{i}=.;
 if netto{i}=1 then do;
  if ct102{i} lt 0 then ct102{i}=0;
  if ct101{i} eq .A or ct101{i} eq .C then ct103{i}=ct102{i};
  else if ct101{i} eq .B then ct103{i}=0;
  else ct103{i}=ct101{i};
  ct104{i}=ct103{i};
 end;
 else ct104{i}=.S;
end;
```

**Variable Name** DIVDY84 – DIVDY04

**Variable Label** Interest, dividend income

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the household income from interest and dividends.

**Method** After 1984 respondents who could not estimate their interest and dividend income directly were asked to select a range from a set of categories. Their choices were:

under 500 DM
500 to 2,000 DM
2,000 to 5,000 DM
5,000 to 10,000 DM
10,000 DM and over

Starting in year 2001 (wave R) an additional item was offered:

10,000 to 20,000 DM
20,000 DM and over

Since year 2002 (wave S) all items are asked for Euro:

under 250 Euro
250 to 1,000 Euro
1,000 to 2,500 Euro
2,500 to 5,000 Euro
5,000 to 10,000 Euro
10,000 Euro and over

These respondents are assigned an interest and dividend amount based on uniformly distributed random numbers within their income range.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.
The original survey variables provided below can be found in the file _H.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984-2004** algorithm:

```
array cx101{*} ah45    bh3801  ch5001  dh5001  eh4101  fh4101  gh4101
               hh4701  ih4401  jh4401  kh4401  lh4401  mh4401  nh4401
               oh4401  ph4401  qh4401  rh4401  sh4401  th4201  uh4201;
                 *** imputed values due to item-non response ***
array cx102{*} xah45   xbh3801 xch5001 xdh5001 xeh4101 xfh4101 xgh4101
               xhh4701 xih4401 xjh4401 xkh4401 xlh4401 xmh4401 xnh4401
               xoh4401 xph4401 xqh4401 xrh4401 xsh4401 xth4201 xuh4201;
array cx103{*} temp184 - temp104
array cx104{*} divdy84 - divdy04

if ah45=.B then ah45=0;
do i = 1 to dim(netto);
cx103{i}=.;
if netto{i}=1 then do;
  if cx101{i} eq .A or cx101{i} eq .C then cx103{i}=cx102{i};
  else cx103{i}=cx101{i}; cx104{i}=cx103{i}; end;
 else cx104{i}=-2;
end;
```

**Variable Name** CHSPT84 – CHSPT04

**Variable Label** Child allowance

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the household income from child allowances.

**Method** In 1984 questions related to this topic were not asked. Child benefits for this year were imputed using information on the number of children in the household and the number of months the benefits were received.
In 1985 to 2000 there was no information regarding the number of months the children allowance was claimed. In all those cases 12 months of claim was supposed.
Since 2001 child allowances is the product of the number of months the children allowance was claimed in the previous year and the average amount per month.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.
The original survey variables provided below can be found in the file _H.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984-2004** algorithm:

```
array dp101{*} out     bh3303  ch4503  dh4503  eh3603  fh3603  gh3603
               hh4503  ih4603  jh4603  kh4603  lh4603  mh4603  nh4603
               oh4603  ph4603  qh50    rh4603  sh4603  th4503  uh4503;
*** imputed values due to item-non response ***
array dp102{*} out     xbh3303 xch4503 xdh4503 xeh3603 xfh3603 xgh3603
               xhh4503 xih4603 xjh4603 xkh4603 xlh4603 xmh4603 xnh4603
               xoh4603 xph4603 xqh50   xrh4603 xsh4603 xth4503 xuh4503;
array dp103{*} kg84    out     out     out     out     out     out
               out     out     out     out     out     out     out
               out     out     out     out     out     out     out;
array dp104{*} temp184 - temp104
array dp106{*} out     out     out     out     out      out     out
               out     out     out     out     out      out     out
               out     out     out     rh4602  sh4602  th4502  uh4502;
array dp105{*} chspt84 - chspt04

do i = 1 to dim(netto);
 dp104{i}=.;
 if netto{i}=1 then do;
  if dp102{i} lt 0 then dp102{i} = 0;
  if dp103{i} lt 0 then dp103{i} = 0;
  if dp106{i} = .A or dp106{i} = .C then dp106{i}=12;
  if dp106{i} in (.B,.)             then dp106{i}=0;

  if dp101{i} = .A or dp101{i} = .C then dp104{i}=dp102{i};
  else if dp101{i} in (.B,.)        then dp104{i}=0;
  else                                   dp104{i}=dp101{i};

       if i=1                  then dp105{i}=dp103{i};          * wave A   *;
  else if (i ge 2 and i le 17) then dp105{i}=dp104{i}*12;       * wave B-Q *;
  else if (i ge 18)            then dp105{i}=dp104{i}*dp106{i}; * wave R ..*;
 end;
 else dp105{i}=-2; end;
```

**Variable Name** HOUSE84 – HOUSE04

**Variable Label** Housing allowance

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the household income from housing allowance.

**Method** Housing allowance is the product of the number of months that benefit was claimed in the previous year and the average amount per month.

**Format** The value of this variable ranges from 0 to 5.000.000.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _H.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm:

```
array dl101{*} ah29    bh2802  ch4002  dh4002  eh3102  fh3102  gh3102
               hh4002  ih4502  jh4502  kh4502  lh4502  mh4502  nh4502
               oh4502  ph4502  qh47    rh4605  sh4605  th4505  uh4505;
array dl102{*} ah30    bh2803  ch4003  dh4003  eh3103  fh3103  gh3103
               hh4003  ih4503  jh4503  kh4503  lh4503  mh4503  nh4503
               oh4503  ph4503  qh48    rh4606  sh4606  th4506  uh4506;
*** imputed values due to item-non response ***
array dl103{*} xah30   xbh2803 xch4003 xdh4003 xeh3103 xfh3103 xgh3103
               xhh4003 xih4503 xjh4503 xkh4503 xlh4503 xmh4503 xnh4503
               xoh4503 xph4503 xqh48   xrh4606 xsh4606 xth4506 xuh4506;
array dl104{*} temp184 - temp104
array dl105{*} temp284 - temp204;
array dl106{*} house84 - house04;

do i = 1 to dim(netto);
 dl104{i}=.;
 dl105{i}=.;
 if dl103{i} lt 0 then dl103{i} = 0;
 if dl101{i} = .A or dl101{i} = .C then dl101{i} = 10;
 if netto{i}=1 then do;
  if dl101{i} eq .B then dl104{i}=0;
  else dl104{i}=dl101{i};
  if dl102{i} eq .A or dl102{i} eq .C then do;
   if dl101{i} le 0 and dl103{i} gt 0 then dl104{i}=12;
   dl105{i}=dl103{i};
  end;
  else if dl102{i} eq .B then dl105{i}=0;
  else dl105{i}=dl102{i};
  dl106{i}=dl104{i}*dl105{i};
 end;
 else dl106{i}=-2;
end;
```

**Variable Name** NURSH84 – NURSH04

**Variable Label** Nursing allowances

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the household income from nursing allowances. Nursing allowances was introduced in the German welfare system in 1996. In 1996-2000 questions related to this topic were only asked for the month of the interview but not for the previous year. Housing allowances for the previous year was imputed using this information. Since 2001 both the numbers of that benefit was claimed in the previous year and the average amount per month were asked.

**Method** Nursing allowances is the product of the number of months that benefit was claimed in the previous year and the average amount per month.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _H.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2000** algorithm: N/a
**2001-2004** algorithm:

```
array dzc101{*} rh4609  sh4609  th4509  uh4509;
*** imputed values due to item-non response ***
array dzc102{*} xrh4609 xsh4609 xth4509 xuh4509;
array dzc103{*} rh4608  sh4608  th4508  uh4508;
array dzc104{*} nursh01 nursh02 nursh03 nursh04;
do i = 1 to dim(dzc101);
 if dzc103{i} eq .A or dzc103{i} = .C then dzc103{i} = 10;
 if dzc101{i} eq .A or dzc101{i} eq .C then do;
  if dzc102{i} gt 0 and dzc103{i} gt 0 then dzc104{i}=dzc102{i}*dzc103{i};
 end;
 else if dzc101{i} eq .B then do;
  dzc104{i}=0;
 end;
 else do;
  dzc104{i}=dzc101{i}*dzc103{i};
 end;
end;
```

**Variable Name** SUBST84 – SUBST04

**Variable Label** Social assistance

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the household income from social assistance.

**Method** Social assistance is the product of the number of months that benefit was claimed in the previous year and the average amount per month. In 1992 through 1994 the amounts of subsistence assistance and special circumstances benefits were not asked. These values have been filled in with imputed values for total social welfare income. In 1995 through 2000 amounts of subsistence assistance and special circumstances benefits are imputed using an algorithm developed by Peter Krause (DIW) based on the benefits received in the present survey month.

**Format** The value of this variable ranges from 0 to 99.999.
This variable is in current year EURO.
The original survey variables provided below can be found in the file _H.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984-2004** algorithm:

```
array dt101{*} ah34     bh3002   ch4202   dh4202   eh3302   fh3302   gh3302
               hh4202   out      out      out      out      out      out
               out      out      out      rh4702   sh4702   th4602   uh4602;
array dt102{*} ah35     bh3003   ch4203   dh4203   eh3303   fh3303   gh3303
               hh4203   out      out      out      out      out      out
               out      out      out      rh4703   sh4703   th4603   uh4603;
*** imputed values due to item-non response ***
array dt103{*} xah35    xbh3003  xch4203  xdh4203  xeh3303  xfh3303  xgh3303
               xhh4203  out      out      out      out      out      out
               out      out      out      xrh4703  xsh4703  xth4603  xuh4603;
*** imputed values due to lacking information in the questionnaire ***
array dt104{*} out      out      out      out      out      out      out
               out      sozye92  sozye93  sozye94  socast95 socast96 socast97
               socast98 socast99 socast00 out      out      out      out;
array dt105{*} temp184 - temp104
array dt106{*} temp284 - temp204;
array dt107{*} subst84 - subst04;

do i = 1 to dim(netto);
 dt105{i}=.; dt106{i}=.;
 if netto{i}=1 then do;
  if dt103{i} lt 0 then dt103{i}=0;
  if dt101{i} eq .B      then dt105{i}=0;
  if dt101{i} in (.A,.C) then dt105{i}=12;
  if dt101{i} ge 0       then dt105{i}=dt101{i};
  if dt102{i} eq .A or dt102{i} eq .C then do;
   if dt101{i} le 0 and dt103{i} gt 0 then dt105{i}=12;
   dt106{i}=dt103{i};  end;
  else if dt102{i} eq .B then dt106{i}=0;
  else dt106{i}=dt102{i};
    if i ge 9 and i le 17 then do;
     if dt104{i}=. then dt104{i}=0;    end;
  if i ge 9 and i le 17 then dt107{i}=dt104{i};
  else dt107{i}=dt105{i}*dt106{i};  end;
 else dt107{i}=-2; end;
```

**Variable Name** SPHLP84 – SPHLP04

**Variable Label** Social assistance for special circumstances

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the household income from Social assistance for special circumstances.

**Method** Social assistance for special circumstances is the product of the number of months that benefit was claimed in the previous year and the average amount per month.
In 1992 through 2000 the amounts of special circumstances benefits were not asked.

**Format** The value of this variable ranges from 0 to 99.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984-2004** algorithm:

```
array dx101{*} ah3601  bh3102  ch4302  dh4302  eh3402  fh3402  gh3402
               hh4302  out     out     out     out     out     out
               out     out     out     rh4705  sh4705  th4605  uh4605;
array dx102{*} ah37    bh3103  ch4303  dh4303  eh3403  fh3403  gh3403
               hh4303  out     out     out     out     out     out
               out     out     out     rh4706  sh4706  th4606  uh4606;
*** imputed values due to item-non response ***
array dx103{*} xah37   xbh3103 xch4303 xdh4303 xeh3403 xfh3403 xgh3403
               xhh4303 out     out     out     out     out     out
               out     out     out     xrh4706 xsh4706 xth4606 xuh4606;
array dx104{*} temp184 - temp104;
array dx105{*} temp284 - temp204;
array dx106{*} sphlp84 - sphlp=4;

do i = 1 to dim(netto);
 dx104{i}=.;
 dx105{i}=.;
 if netto{i}=1 then do;
  if dx103{i} lt 0 then dx103{i}=0;
  if dx101{i} eq .B      then dx104{i}=0;
  if dx101{i} in (.A,.C) then dx104{i}=10;
  if dx101{i} ge 0       then dx104{i}=dx101{i};

  if dx102{i} eq .A or dx102{i} eq .C then do;
   if dx101{i} le 0 and dx103{i} gt 0 then dx104{i}=12;
   dx105{i}=dx103{i};
  end;
  else if dx102{i} eq .B then dx105{i}=0;
  else dx105{i}=dx102{i};
  dx106{i}=dx104{i}*dx105{i};
 end;
 else dx106{i}=-2;
end;
```

**Variable Name** HSUP84 – HSUP04

**Variable Label** Housing support for owner-occupiers

**Unit of Observation** Household

**Period** Annual

**Description** This variable represents the household income from direct housing support for owner-occupiers.

**Method** In 1996 German law established direct housing subsidy payments for owner-occupiers. Starting in the 2000 survey a separate question was asked about income from this source. Direct housing subsidy payments for respondents who bought homes between 1996 and 1999 were imputed using information about the year of construction, acquisition of ownership and number of children in the household.

**Format** The value of this variable ranges from 0 to 9.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _H.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm:

```
array bzc101{*} misses  misses  misses  misses  misses  misses  misses
                misses  misses  misses  misses  misses  ms3904  ns3904
                os3904  ps3904  qh3904  rh3904  sh3904  th3504  uh3504;
*** imputed values due to item-non response ***
array bzc102{*} misses  misses  misses  misses  misses  misses  misses
                misses  misses  misses  misses  misses  xms3904 xns3904
                xos3904 xps3904 xqh3904 xrh3904 xsh3904 xth3504 xuh3504;
array bzc103{*} temp184 - temp104
array bzc106{*} hsup84 - hsup04;

do i = 1 to dim(netto);
 bzc103{i}=.;
 if netto{i}=1 then do;
  if bzc102{i} lt 0 then bzc102{i} = 0;
  if bzc101{i} eq .A or bzc101{i} eq .C then bzc103{i}=bzc102{i};
  else if bzc101{i} eq .B               then bzc103{i}=0;
  else                                       bzc103{i}=bzc101{i};
  bzc106{i}=bzc103{i};
 end;
 else bzc106{i}=-2;
end;
```

| | |
|---|---|
| **Variable Name** | FRENTY84 – FRENTY04 |
| **Variable Label** | Impute Income from rental and leasing |
| **Unit of Observation** | Household |
| **Description** | This variable indicates if income from rental and leasing (RENTY$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FOPERY84 – FOPERY04 |
| **Variable Label** | Impute Operation, maintenance costs |
| **Unit of Observation** | Household |
| **Description** | This variable indicates if operation or maintenance costs (OPERY$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FDIVDY84 – FDIVDY04 |
| **Variable Label** | Impute Interest, dividend income |
| **Unit of Observation** | Household |
| **Description** | This variable indicates if income from interest or dividends (DIVDY$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FCHSPT84 – FCHSPT04 |
| **Variable Label** | Impute Child allowance |
| **Unit of Observation** | Household |
| **Description** | This variable indicates if income from child allowances (CHSPT$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FHOUSE84 – FHOUSE04 |
| **Variable Label** | Impute Housing benefit |
| **Unit of Observation** | Household |
| **Description** | This variable indicates if income from housing benefit (HOUSE$$) has been imputed. The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FNURSH84 – FNURSH04 |
| **Variable Label** | Impute Nursing allowances |
| **Unit of Observation** | Household |
| **Description** | This variable indicates if income from nursing allowances (NURSH$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FSUBST84 – FSUBST04 |
| **Variable Label** | Impute Social assistance |
| **Unit of Observation** | Household |
| **Description** | This variable indicates if income from social assistance (SUBST$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FSPHLP84 – FSPHLP04 |
| **Variable Label** | Impute Social assistance for special circumstances |
| **Unit of Observation** | Household |
| **Description** | This variable indicates if income from social assistance for special circumstances (SPHLP$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FHSUP84 – FHSUP04 |
| **Variable Label** | Impute Housing support for owner-occupiers |
| **Unit of Observation** | Household |
| **Description** | This variable indicates if income from housing support for owner-occupiers (HSUP$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | IJOB184 – IJOB104 |
| **Variable Label** | Wages, Salary from main job |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents wages or salary from main job of individuals in the household 16 years of age and older. |
| **Method** | Wages or salary from main job is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

| | | |
|---|---|---|
| **1984-2004** algorithm: | IJOB1$$ = ($P2A02 * $P2A03) | ($$=84 – 04, $ = A, B, … , U) |

**Variable Name** IJOB284 – IJOB204

**Variable Label** Income from secondary employment

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from secondary employment of individuals in the household 16 years of age and older.

**Method** Income from secondary employment is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: IJOB2$$ = ($P2C02 * $P2C03) ($$=84 – 04, $ = A, B, … , U)

| | |
|---|---|
| **Variable Name** | ISELF84 – ISELF04 |
| **Variable Label** | Income from self-employment |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from self-employment of individuals in the household 16 years of age and older. |
| **Method** | Income from self-employment is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

**1984-2004** algorithm: ISELF$$ = ($P2B02 * $P2B03) ($$=84 – 04, $ = A, B, … , U)

| | |
|---|---|
| **Variable Name** | IOLDY84 – IOLDY04 |
| **Variable Label** | Combined old-age, disability and civil servants pensions |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from combined old-age, disability and civil servants pensions of individuals in the household 16 years of age and older.<br>In 2002 and 2003 separate questions regarding income from private or company pension were asked. Thus these incomes components are not included in old-age, disability and civil servants pensions in the those years. |
| **Method** | Income from combined old-age, disability and civil servants pensions is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

| | | |
|---|---|---|
| **1984-2004** algorithm: | IOLDY$$ = ($P2D02 * $P2D03) | ($$=84 – 04, $ = A, B, … , U) |

| | |
|---|---|
| **Variable Name** | IWIDY84 – IWIDY04 |
| **Variable Label** | Combined widows and orphans pension |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from combined widows and orphans pension of individuals in the household 16 years of age and older.<br>In 2002 and 2003 separate questions regarding income from private or company pension were asked. Thus these incomes components are not included in widows and orphans pension in the those years. |
| **Method** | Income from combined widows and orphans pension is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

| | | |
|---|---|---|
| **1984-2004** algorithm: | IWIDY$$ = ($P2E02 * $P2E03) | ($$=84 – 04, $ = A, B, … , U) |

| | |
|---|---|
| **Variable Name** | ICOMP84 – ICOMP04 |
| **Variable Label** | Combined company pension (surviving dependants c.p.) |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from combined company pension of individuals in the household 16 years of age and older. In 1984-2001 and since 2004 specific questions related to this topic were not asked. Thus these income component is included in old-age, disability and civil servants pensions (IOLDY$$) in the those years. |
| **Method** | Income from combined company pension is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

| | | |
|---|---|---|
| **1984-2001** algorithm: | N/a | |
| **2002-2003** algorithm: | ICOMP$$ = ($P2P02 * $P2P03) | ($$=02 – 03, $ = S-T) |
| **2004** algorithm: | N/a | |

| | |
|---|---|
| **Variable Name** | IPRVP84 – IPRVP04 |
| **Variable Label** | Combined private pension (old-age, accident, disability) |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from combined private pension of individuals in the household 16 years of age and older. In 1984-2001 and since 2004 specific questions related to this topic were not asked. Thus these income component is included in old-age, disability and civil servants pensions (IOLDY$$) in the those years. |
| **Method** | Income from combined private pension is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

| | | |
|---|---|---|
| **1984-2001** algorithm: | N/a | |
| **2002-2003** algorithm: | IPRVP$$ = ($P2Q02 * $P2Q03) | ($$=02 – 03, $ = S-T) |
| **2004** algorithm: | N/a | |

| | |
|---|---|
| **Variable Name** | IUNBY84 – IUNBY04 |
| **Variable Label** | Unemployment benefit |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from unemployment benefit of individuals in the household 16 years of age and older. |
| **Method** | Income from unemployment benefit is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

**1984-2004** algorithm: IUNBY$$ = ($P2F02 * $P2F03) ($$=84 – 04, $ = A, B, … , U)

| | |
|---|---|
| **Variable Name** | IUNAY84 – IUNAY04 |
| **Variable Label** | Unemployment assistance |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from unemployment assistance of individuals in the household 16 years of age and older. |
| **Method** | Income from unemployment assistance is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

**1984-2004** algorithm: IUNAY$$ = ($P2G02 * $P2G03) ($$=84 – 04, $ = A, B, … , U)

| | |
|---|---|
| **Variable Name** | ISUBY84 – ISUBY04 |
| **Variable Label** | Subsistence allowance |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from subsistence allowance of individuals in the household 16 years of age and older. |
| **Method** | Income from subsistence allowance is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

**1984-2004** algorithm: ISUBY$$ = ($P2H02 * $P2H03) ($$=84 – 04, $ = A, B, … , U)

| | |
|---|---|
| **Variable Name** | IERET84 – IERET04 |
| **Variable Label** | Old-age transition benefit |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from old-age transition benefit of individuals in the household 16 years of age and older. In 1984-1995 and since 2002 questions related to this topic were not asked. |
| **Method** | Income from old-age transition benefit is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

| | | |
|---|---|---|
| **1984-1995** algorithm: | N/a | |
| **1996-2001** algorithm: | IERET$$ = ($P2I02 * $P2I03) | ($$=96 – 01, $ = M, N, … , R) |
| **2002-2004** algorithm: | N/a | |

| | |
|---|---|
| **Variable Name** | IMATY84 – IMATY04 |
| **Variable Label** | Maternity benefit |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from maternity benefit of individuals in the household 16 years of age and older. |
| **Method** | Income from maternity benefit is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

**1984-2004** algorithm: IMATY$$ = ($P2J02 * $P2J03) ($$=84 – 04, $ = A, B, … , U)

| | |
|---|---|
| **Variable Name** | ISTUY84 – ISTUY04 |
| **Variable Label** | Student grants |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from student grants of individuals in the household 16 years of age and older. |
| **Method** | Income from student grants is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

**1984-2004** algorithm: ISTUY$$ = ($P2K02 * $P2K03) ($$=84 – 04, $ = A, B, … , U)

| | |
|---|---|
| **Variable Name** | IMILT84 – IMILT04 |
| **Variable Label** | Military community service pay |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from military community service pay of individuals in the household 16 years of age and older. In 1984-1995 questions related to this topic were not asked. |
| **Method** | Income from military community service pay is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

| | | |
|---|---|---|
| **1984-1995** algorithm: | N/a | |
| **1996-2004** algorithm: | IMILT$$ = ($P2L02 * $P2L03) | ($$=96 – 04, $ = M, N, … , U) |

**Variable Name** IALIM84 – IALIM04

**Variable Label** Alimony

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from alimony of individuals in the household 16 years of age and older. In 1984-2000 specific questions related to this topic were not asked. Alimony is included in private transfers received (IELSE$$) in the those years.

**Method** Income from alimony is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2000** algorithm: N/a
**2001-2004** algorithm: IALIM$$ = ($P2O02 * $P2O03) ($$=00 – 04, $ = Q, R, … , U)

| | |
|---|---|
| **Variable Name** | IELSE84 – IELSE04 |
| **Variable Label** | Private Transfers received |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from private transfers of individuals in the household 16 years of age and older. In 1984-2000 alimony is included in private transfers. Since 2001 a specific question regarding alimony (IALIM$$) were asked, thus alimony is no longer included in private transfers received. |
| **Method** | Income from private transfers is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

**1984-2004** algorithm: IELSE$$ = ($P2M02 * $P2M03) ($$=84 – 04, $ = A, B, … , U)

| | |
|---|---|
| **Variable Name** | I13LY84 – I13LY04 |
| **Variable Label** | 13th monthly salary |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from 13th monthly salary of individuals in the household 16 years of age and older. |
| **Method** | Transcribed variable. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _P.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

**1984-2004** algorithm: I13LY$$ = Y13 ($$=84 – 04)

13th monthly salary variable list by survey year - each entry denoted in algorithm as Y13:

ap3902 bp5902 cp5902 dp5902 ep5402 fp7202 gp7202 hp6702 ip6702 jp7702 kp7702 lp8202 mp6802 np6802 op5902 pp7702 qp7702 rp7702 sp7702 tp9502 up8002

**Variable Name** I14LY84 – I14LY04

**Variable Label** 14th monthly salary

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from 14th monthly salary of individuals in the household 16 years of age and older.

**Method** Transcribed variable.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: I14LY$$ = Y14 ($$=84 – 04)

14th monthly salary variable list by survey year - each entry denoted in algorithm as Y14:

ap3904 bp5904 cp5904 dp5904 ep5404 fp7204 gp7204 hp6704 ip6704 jp7704 kp7704 lp8204 mp6804 np6804 op5904 pp7704 qp7704 rp7704 sp7704 tp9504 up8004

| | |
|---|---|
| **Variable Name** | IXMAS84 – IXMAS04 |
| **Variable Label** | Christmas bonus |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from Christmas bonus of individuals in the household 16 years of age and older. |
| **Method** | Transcribed variable. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _P.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

**1984-2004** algorithm: IXMAS$$ = YXMS ($$=84 – 04)

Christmas bonus variable list by survey year - each entry denoted in algorithm as YXMS:

ap3906 bp5906 cp5906 dp5906 ep5406 fp7206 gp7206 hp6706 ip6706 jp7706 kp7706 lp8206 mp6806 np6806 op5906 pp7706 qp7706 rp7706 sp7706 tp9506 up8006

| | |
|---|---|
| **Variable Name** | IHOLY84 – IHOLY04 |
| **Variable Label** | Vacation bonus |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from vacation bonus of individuals in the household 16 years of age and older. |
| **Method** | Transcribed variable. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _P.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

**1984-2004** algorithm: IHOLY$$ = YHOL ($$=84 – 04)

Vacation bonus variable list by survey year - each entry denoted in algorithm as YHOL:

ap3908 bp5908 cp5908 dp5908 ep5408 fp7208 gp7208 hp6708 ip6708 jp7708 kp7708 lp8208 mp6808 np6808 op5908 pp7708 qp7708 rp7708 sp7708 tp9508 up8008

| | |
|---|---|
| **Variable Name** | IGRAY84 – IGRAY04 |
| **Variable Label** | Profit-sharing |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from profit sharing of individuals in the household 16 years of age and older. |
| **Method** | Transcribed variable. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _P.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

**1984-2004** algorithm: IGRAY$$ = YGRY ($$=84 – 04)

Profit sharing variable list by survey year - each entry denoted in algorithm as YGRY:

ap3910 bp5910 cp5910 dp5910 ep5410 fp7210 gp7210 hp6710 ip6710 jp7710 kp7710 lp8210 mp6810 np6810 op5910 pp7710 qp7710 rp7710 sp7710 tp9510 up8010

**Variable Name** IOTHY84 – IOTHY04

**Variable Label** Other bonuses

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from other bonuses of individuals in the household 16 years of age and older. In 1984 questions related to this topic were not asked. They are included in income from profit sharing (IGRAY84).

**Method** Transcribed variable.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984** algorithm: N/a
**1985-2004** algorithm: IOTHY$$ = YOTH ($$=85 – 04)

Other bonuses variable list by survey year - each entry denoted in algorithm as YOTH:

bp5912 cp5912 dp5912 ep5412 fp7212 gp7212 hp6712 ip6712 jp7712 kp7712 lp8212 mp6812 np6812 op5912 pp7712 qp7712 rp7712 sp7712 tp9512 up8012

| | |
|---|---|
| **Variable Name** | IGRV184 – IGRV104 |
| **Variable Label** | Statutory pension insurance |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from statutory pension insurance of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked. Since 2002 the statutory pension insurance did also include the social miners insurance pension (ISMP1$$) and farmers pension (IAGR1$$). |
| **Method** | Income from statutory pension insurance is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _P and _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

| | |
|---|---|
| **1984-1985** algorithm: | N/a |
| **1986-2004** algorithm: | IGRV1$$ = ($P2D02 * YSPI) ($$=86 – 04, $ = C, D, … , U) |

Statutory pension insurance variable list by survey year - each entry denoted in algorithm as YSPI:

cp6101 dp6101 ep5601 fp7401 gp7401 hp6901 ip6901 jp7901 kp7901 lp8401 mp7001 np7001 op6101 pp7901 qp7901 rp7901 tp9701 up8201

**Variable Name** IGRV284 – IGRV204

**Variable Label** Widows and orphans statutory pension insurance

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from widows and orphans statutory pension insurance of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked. Since 2002 the widows and orphans statutory pension insurance does also include the widows and orphans social miners insurance pension (ISMP2$$) and widows and orphans farmers pension (IAGR2$$).

**Method** Income from widows and orphans statutory pension insurance is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2004** algorithm: IGRV2$$ = ($P2E02 * YWID) ($$=86 – 04, $ = C, D, … , U)

Widows and orphans statutory pension insurance variable list by survey year - each entry denoted in algorithm as YWID:

cp6110 dp6110 ep5610 fp7410 gp7410 hp6910 ip6910 jp7911 kp7911 lp8411 mp7010 np7010 op6110 pp7910 qp7910 rp7910 tp9702 up8202;

| | |
|---|---|
| **Variable Name** | ISMP184 – ISMP104 |
| **Variable Label** | Social miners insurance pension |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from social miners insurance pension of individuals in the household 16 years of age and older. In 1984-1985 and since 2002 specific questions related to this topic were not asked. Since 2002 this income component is included in the statutory pension insurance (IGRV1$$). |
| **Method** | Income from social miners insurance pension is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _P and _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

| | | |
|---|---|---|
| **1984-1985** algorithm: | N/a | |
| **1986-2001** algorithm: | ISMP1$$ = ($P2D02 * YSMP) | ($$=86 – 01, $ = C, D, … , R) |
| **2002-2004** algorithm: | N/a | |

Social miners insurance pension variable list by survey year - each entry denoted in algorithm as YSMP:

cp6102 dp6102 ep5602 fp7402 gp7402 hp6902 ip6902 jp7903 kp7903 lp8403 mp7002 np7002 op6102 pp7902 qp7902 rp7902

**Variable Name** ICIV184 – ICIV104

**Variable Label** Civil servant pension

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from civil servant pension of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked.

**Method** Income from civil servant pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2004** algorithm: ICIV1$$ = ($P2D02 * YCIV) ($$=86 – 04, $ = C, D, … , U)

Civil servant pension variable list by survey year - each entry denoted in algorithm as YCIV:

cp6103 dp6103 ep5603 fp7403 gp7403 hp6903 ip6903 jp7904 kp7904 lp8404 mp7003 np7003 op6103 pp7903 qp7903 rp7903 tp9703 up8203 ;

**Variable Name** IWAR184 – IWAR104

**Variable Label** War victim pension

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from war victim pension of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked.

**Method** Income from war victim pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2004** algorithm: IWAR1$$ = ($P2D02 * YWAR) ($$=86 – 04, $ = C, D, … , U)

War victim pension variable list by survey year - each entry denoted in algorithm as YWAR:

cp6104 dp6104 ep5604 fp7404 gp7404 hp6904 ip6904 jp7905 kp7905 lp8405 mp7004 np7004 op6104 pp7904 qp7904 rp7904 tp9705 up8205

**Variable Name** IAGR184 – IAGR104

**Variable Label** Farmer Pension

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from farmer pension of individuals in the household 16 years of age and older. In 1984-1985 and since 2002 specific questions related to this topic were not asked. Since 2002 this income component is included in the statutory pension insurance (IGRV1$$).

**Method** Income from farmer pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2001** algorithm: IAGR1$$ = ($P2D02 * YAGR) ($$=86 – 01, $ = C, D, … , R)
**2001-2004** algorithm: N/a

Farmer pension variable list by survey year - each entry denoted in algorithm as YAGR:

cp6105 dp6105 ep5605 fp7405 gp7405 hp6905 ip6905 jp7906 kp7906 lp8406 mp7005 np7005 op6105 pp7905 qp7905 rp7905

**Variable Name** IGUV184 – IGUV104

**Variable Label** Statutory accident insurance pension

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from statutory accident insurance pension of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked.

**Method** Income from statutory accident insurance pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2004** algorithm: IGUV1$$ = ($P2D02 * YGUV) ($$=86 – 04, $ = C, D, … , U)

Statutory accident insurance pension variable list by survey year - each entry denoted in algorithm as YGUV:

cp6106 dp6106 ep5606 fp7406 gp7406 hp6906 ip6906 jp7907 kp7907 lp8407 mp7006 np7006 op6106 pp7906 qp7906 rp7906 tp9707 up8207

**Variable Name** IVBL184 – IVBL104

**Variable Label** Supplementary benefits for civil servants

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from supplementary benefits for civil servants of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked.

**Method** Income from supplementary benefits for civil servants is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2004** algorithm: IVBL1$$ = ($P2D02 * YVBL) ($$=86 – 04, $ = C, D, … , U)

Supplementary benefits for civil servants variable list by survey year - each entry denoted in algorithm as YVBL:

cp6107 dp6107 ep5607 fp7407 gp7407 hp6907 ip6907 jp7908 kp7908 lp8408 mp7007 np7007 op6107 pp7907 qp7907 rp7907 tp9709 up8209

**Variable Name** ICOM184 – ICOM104

**Variable Label** Company pension

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from company pension of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked.

**Method** Income from company pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2004** algorithm: ICOM1$$ = ($P2D02 * YCOM) ($$=86 – 04, $ = C, D, … , U)

Company pension variable list by survey year - each entry denoted in algorithm as YCOM:

cp6108 dp6108 ep5608 fp7408 gp7408 hp6908 ip6908 jp7909 kp7909 lp8409 mp7008 np7008 op6108 pp7908 qp7908 rp7908 tp9711 up8211

**Variable Name** IPRV184 – IPRV104

**Variable Label** Private pension

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from private pension of individuals in the household 16 years of age and older. In 1984-2002 specific questions related to this topic were not asked.

**Method** Income from private pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2002** algorithm: N/a
**2003-2004** algorithm: IPRV1$$ = ($P2D02 * YPRV) ($$=03 – 04, $ = T, … , U)

Private pension variable list by survey year - each entry denoted in algorithm as YPRV:

tp9713 up8213

| | |
|---|---|
| **Variable Name** | ISON184 – ISON104 |
| **Variable Label** | Other pension |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from other pension of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked. |
| **Method** | Income from other pension is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _P and _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

| | | |
|---|---|---|
| **1984-1985** algorithm: | N/a | |
| **1986-2004** algorithm: | ISON1$$ = ($P2D02 * YSON) | ($$=86 – 04, $ = C, D, … , U) |

Other pension variable list by survey year - each entry denoted in algorithm as YSON:

cp6109 dp6109 ep5609 fp7409 gp7409 hp6909 ip6909 jp7910 kp7910 lp8410 mp7009 np7009 op6109 pp7909 qp7909 rp7909 tp9715 up8215

**Variable Name** ISMP284 – ISMP204

**Variable Label** Widows and orphans social miners insurance pension

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from widows and orphans social miners insurance pension of individuals in the household 16 years of age and older. In 1984-1985 and since 2002 specific questions related to this topic were not asked. Since 2002 this income component is included in the statutory pension insurance (IGRV2$$).

**Method** Income from widows and orphans social miners insurance pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2001** algorithm: ISMP2$$ = ($P2E02 * YSMP) ($$=86 – 01, $ = C, D, … , R)
**2002-2004** algorithm: N/a

Widows and orphans social miners insurance pension variable list by survey year - each entry denoted in algorithm as YSMP:

cp6111 dp6111 ep5611 fp7411 gp7411 hp6911 ip6911 jp7913 kp7913 lp8413 mp7011 np7011 op6111 pp7911 qp7911 rp7911

| | |
|---|---|
| **Variable Name** | ICIV284 – ICIV204 |
| **Variable Label** | Widows and orphans civil servant pension |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from widows and orphans civil servant pension of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked. |
| **Method** | Income from widows and orphans civil servant pension is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _P and _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

| | | |
|---|---|---|
| **1984-1985** algorithm: | N/a | |
| **1986-2004** algorithm: | ICIV2$$ = ($P2E02 * YCIV) | ($$=86 – 04, $ = C, D, … , U) |

Widows and orphans civil servant pension variable list by survey year - each entry denoted in algorithm as YCIV:

cp6112 dp6112 ep5612 fp7412 gp7412 hp6912 ip6912 jp7914 kp7914 lp8414 mp7012 np7012 op6112 pp7912 qp7912 rp7912 tp9704 up8204;

**Variable Name** IWAR284 – IWAR204

**Variable Label** Widows and orphans war victim pension

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from widows and orphans war victim pension of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked.

**Method** Income from widows and orphans war victim pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2004** algorithm: IWAR2$$ = ($P2E02 * YWAR) ($$=86 – 04, $ = C, D, … , U)

Widows and orphans war victim pension variable list by survey year - each entry denoted in algorithm as YWAR:

cp6113 dp6113 ep5613 fp7413 gp7413 hp6913 ip6913 jp7915 kp7915 lp8415 mp7013 np7013 op6113 pp7913 qp7913 rp7913 tp9706 up8206

**Variable Name** IAGR284 – IAGR204

**Variable Label** Widows and orphans farmer Pension

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from widows and orphans farmer pension of individuals in the household 16 years of age and older. In 1984-1985 and since 2002 specific questions related to this topic were not asked. Since 2002 this income component is included in the statutory pension insurance (IGRV2$$).

**Method** Income from widows and orphans farmer pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2001** algorithm: IAGR2$$ = ($P2E02 * YAGR) ($$=86 – 01, $ = C, D, … , R)
**2001-2004** algorithm: N/a

Widows and orphans Farmer pension variable list by survey year - each entry denoted in algorithm as YAGR:

cp6114 dp6114 ep5614 fp7414 gp7414 hp6914 ip6914 jp7916 kp7916 lp8416 mp7014 np7014 op6114 pp7914 qp7914 rp7914

**Variable Name** IGUV284 – IGUV204

**Variable Label** Widows and orphans statutory accident insurance

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from widows and orphans statutory accident insurance pension of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked.

**Method** Income from widows and orphans statutory accident insurance pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2004** algorithm: IGUV2$$ = ($P2E02 * YGUV) ($$=86 – 04, $ = C, D, … , U)

Widows and orphans statutory accident insurance pension variable list by survey year - each entry denoted in algorithm as YGUV:

cp6115 dp6115 ep5615 fp7415 gp7415 hp6915 ip6915 jp7917 kp7917 lp8417 mp7015 np7015 op6115 pp7915 qp7915 rp7915 tp9708 up8208

| | |
|---|---|
| **Variable Name** | IVBL284 – IVBL204 |
| **Variable Label** | Widows and orphans supplement. benefits for civil servants |
| **Unit of Observation** | Individual |
| **Period** | Annual |
| **Description** | This variable represents income from widows and orphans supplementary benefits for civil servants of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked. |
| **Method** | Income from widows and orphans supplementary benefits for civil servants is the product of the number of months that income was received in the previous year and the average amount per month.<br>If the information about the number of months is missing, the sample mean of that variable has been assigned. |
| **Format** | The value of this variable ranges from 0 to 999.999.<br>This variable is in current year EURO.<br><br>The original survey variables provided below can be found in the file _P and _PKAL.<br>This algorithm omits individuals with survey non-responses. |

**Algorithm**

| | |
|---|---|
| **1984-1985** algorithm: | N/a |
| **1986-2004** algorithm: | IVBL2$$ = ($P2E02 * YVBL) ($$=86 – 04, $ = C, D, … , U) |

Widows and orphans supplementary benefits for civil servants variable list by survey year - each entry denoted in algorithm as YVBL:

cp6116 dp6116 ep5616 fp7416 gp7416 hp6916 ip6916 jp7918 kp7918 lp8418 mp7016 np7016 op6116 pp7916 qp7916 rp7916 tp9710 up8210

**Variable Name** ICOM284 – ICOM204

**Variable Label** Widows and orphans company pension

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from widows and orphans company pension of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked.

**Method** Income from widows and orphans company pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2004** algorithm: ICOM2$$ = ($P2E02 * YCOM) ($$=86 – 04, $ = C, D, … , U)

Widows and orphans company pension variable list by survey year - each entry denoted in algorithm as YCOM:

cp6117 dp6117 ep5617 fp7417 gp7417 hp6917 ip6917 jp7919 kp7919 lp8419 mp7017 np7017 op6117 pp7917 qp7917 rp7917 tp9712 up8212

**Variable Name** IPRV284 – IPRV204

**Variable Label** Widows and orphans private pension

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from widows and orphans private pension of individuals in the household 16 years of age and older. In 1984-2002 specific questions related to this topic were not asked.

**Method** Income from widows and orphans private pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2002** algorithm: N/a
**2003-2004** algorithm: IPRV2$$ = ($P2E02 * YPRV) ($$=03 – 04, $ = T, … , U)

Widows and orphans private pension variable list by survey year - each entry denoted in algorithm as YPRV:

tp9714 up8214

**Variable Name** ISON284 – ISON204

**Variable Label** Other widows or orphans pension

**Unit of Observation** Individual

**Period** Annual

**Description** This variable represents income from other widows or orphans pension of individuals in the household 16 years of age and older. In 1984-1985 specific questions related to this topic were not asked.

**Method** Income from other widows or orphans pension is the product of the number of months that income was received in the previous year and the average amount per month.
If the information about the number of months is missing, the sample mean of that variable has been assigned.

**Format** The value of this variable ranges from 0 to 999.999.
This variable is in current year EURO.

The original survey variables provided below can be found in the file _P and _PKAL.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1985** algorithm: N/a
**1986-2004** algorithm: ISON2$$ = ($P2E02 * YSON) ($$=86 – 04, $ = C, D, … , U)

Other widows or orphans pension variable list by survey year - each entry denoted in algorithm as YSON:

cp6118 dp6118 ep5618 fp7418 gp7418 hp6918 ip6918 jp7920 kp7920 lp8420 mp7018 np7018 op6118 pp7918 qp7918 rp7918 tp9716 up8216

| | |
|---|---|
| **Variable Name** | FJOB184 – FJOB104 |
| **Variable Label** | Imputation flag: Wages, Salary from main job |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from wages or salary from main job (IJOB1$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FJOB284 – FJOB204 |
| **Variable Label** | Imputation flag: Income from secondary job |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from secondary job (IJOB2$$) has been imputed. The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FSELF84 – FSELF04 |
| **Variable Label** | Imputation flag: Income from self-employment |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from self-employment (ISELF$$) has been imputed. The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FOLDY84 – FOLDY04 |
| **Variable Label** | Imputation flag: combined old-age, civil servants pensions. |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from combined old-age, civil servants pensions (IOLDY$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FWIDY84 – FWIDY04 |
| **Variable Label** | Imputation flag: combined widows / orphans pension |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from combined widows or orphans pension (IWIDY$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FCOMP84 – FCOMP04 |
| **Variable Label** | Imputation flag: Combined company pension |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from combined company pension (ICOMP$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FPRVP84 – FPRVP04 |
| **Variable Label** | Imputation flag: Combined private pension |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from combined private pension (IPRVP$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FUNBY84 – FUNBY04 |
| **Variable Label** | Imputation flag: Unemployment benefit |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from unemployment benefit (IUNBY$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FUNAY84 – FUNAY04 |
| **Variable Label** | Imputation flag: Unemployment assistance |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from unemployment assistance (IUNAY$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FSUBY84 – FSUBY04 |
| **Variable Label** | Imputation flag: Subsistence allowance |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from subsistence allowance (ISUBY$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FERET84 – FERET04 |
| **Variable Label** | Imputation flag: Old-age transition benefit |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from old-age transition benefit (IERET$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FMATY84 – FMATY04 |
| **Variable Label** | Imputation flag: Maternity benefit |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from maternity benefit (HSUP$$) has been imputed. The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FSTUY84 – FSTUY04 |
| **Variable Label** | Imputation flag: Student grants |
| Unit of Observation | Individual |
| **Description** | This variable indicates if income from student grants (ISTUY$$) has been imputed. The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FMILT84 – FMILT04 |
| **Variable Label** | Imputation flag: Military / community service pay |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from military or community service pay (IMILT$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FALIM84 – FALIM04 |
| **Variable Label** | Imputation flag: Alimony |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from alimony (IALIM$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FELSE84 – FELSE04 |
| **Variable Label** | Imputation flag: Private Transfers received |
| Unit of Observation | Individual |
| **Description** | This variable indicates if income from private transfers received (IELSE$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | F13LY84 – F13LY04 |
| **Variable Label** | Imputation flag: 13th monthly salary |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from 13th monthly salary (I13LY$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | F14LY84 – F14LY04 |
| **Variable Label** | Imputation flag: 14th monthly salary |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from 14th monthly salary (I14LY$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FXMAS84 – FXMAS04 |
| **Variable Label** | Imputation flag: Christmas bonus |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from Christmas bonus (IXMAS$$) has been imputed.<br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FHOLY84 – FHOLY04 |
| **Variable Label** | Imputation flag: Vacation bonus |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from Vacation bonus (IHOLY$$) has been imputed. The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FGRAY84 – FGRAY04 |
| **Variable Label** | Imputation flag: Profit-sharing / Gratifications |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from profit-sharing (IGRAY$$) has been imputed. The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FOTHY84 – FOTHY04 |
| **Variable Label** | Imputation flag: Other bonuses |
| Unit of Observation | Individual |
| **Description** | This variable indicates if income from other bonuses (IOTHY$$) has been imputed. The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FGRV184 – FGRV104 |
| **Variable Label** | Imputation flag: statutory pension insurance |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from statutory pension insurance (IGRV1$$) has been imputed. In order to reduce complexity and because of minor incidence, the imputation flag variable FGVR1 encompasses also imputation of item-non-response on any of the considered variables related to (own) pension income.<br><br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

| | |
|---|---|
| **Variable Name** | FGRV284 – FGRV204 |
| **Variable Label** | Imputation flag: widows or orphans statutory pension insurance |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates if income from widows or orphans statutory pension insurance (IGRV2$$) has been imputed. In order to reduce complexity and because of minor incidence, the imputation flag variable FGVR2 encompasses also imputation of item-non-response on any of the considered variables related to Widows and orphans pension income.<br><br>The predominant imputation technique used to fill in missing values is based on the row and column imputation procedure developed by Little and Su (1989). In the case of lacking longitudinal data purely cross-sectional imputation techniques are applied. For further details, see: Grabka and Frick (2003). |
| **Method** | In the original SOEP data there are three types of missing values. These missing values can be interpreted as:<br><br>-1 = no answer or do not know<br>-2 = does not apply<br>-3 = original value was deleted because it was found to be implausible<br><br>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only. |
| **Format** | 0 = Not Imputed<br>1 = Fully Imputed<br><br>This algorithm omits individuals with survey non-responses. |

**Variable Name** M1110184 - M1110104

**Variable Label** Whether spent night in hospital in last year

**Unit of Observation** Individual

**Description** Indicates whether person stayed overnight in a hospital at any time in previous year
In 1990 and 1993 this information was not asked.

**Method** Transcribed variable.

**Format** -2 = N/A - Child
-1 = Item non-response
0 = Did not stay overnight in a hospital
1 = Stayed overnight in a hospital

The original survey variables provided below can be found in the file _P.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984-2004** algorithm:

```
array hosp{*}   AP5101     BP7201     CP7201     DP7201     EP6901      FP8501      null          HP8001
        IP8001       null         KP8701     LP9301     MP7901     NP8301     OP7201     PP100
       QP99         RP99        SP98        TP101       up94;
array nhosp{*} M1110184 M1110185 M1110186 M1110187 M1110188 M1110189 M1110190 M1110191
         M1110192 M1110193 M1110194 M1110195 M1110196 M1110197 M1110198 M1110199
         M1110100 M1110101 M1110102 M1110103 M1110104;
do i = 1 to dim(netto);
 if netto{i} in (1,5) then do;
  if i=1 and hosp{i}=-2 then nhosp{i}=0;
  if hosp{i} in (1,2)   then nhosp{i}=2-hosp{i};
  if hosp{i} in (-1,-3) then nhosp{i}=.-1;
  if hosp{i} in (-2)    then nhosp{i}=-2;
 end;
 else nhosp{i}=-2;
end;
```

**Variable Name** M1110284 - M1110204

**Variable Label** Number of nights (days) stayed in hospital overnight in past year

**Unit of Observation** Individual

**Description** Number of nights (days) person stayed overnight in a hospital in previous year
In 1990 and 1993 this information was not asked.

**Method** Transcribed variable.

**Format** -2 = N/A - Child
-1 = Item non-response
0 = Did not stay overnight in a hospital
1-365 = Nights (days) spent in hospital

The original survey variables provided below can be found in the file _P.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984-2004** algorithm:

```
array hospd{*} ap5103    bpP7203   CP7203    DP7203    EP6903    FP8503    null      HP8003
               IP8003    null      kp8703    lp9303    mp7903    NP8303    OP7203    PP10102
               QP10002   RP10001   sp9901    tpP10201  up9501;
array nhospd{*} M1110284 M1110285 M1110286 M1110287 M1110288 M1110289 M1110290 M1110291
               M1110292 M1110293 M1110294 M1110295 M1110296 M1110297 M1110298 M1110299
               M1110200 M1110201 M1110202 M1110203 M1110204;
do i = 1 to dim(netto);
 if netto{i} in (1,5) then do;
  if hospd{i}>=0 then nhospd{i}=hospd{i};
  if hospd{i} in (-1,-3) then nhospd{i}=.M;
  if hospd{i} in (-2)    then nhospd{i}= 0;
 end;
 else nhospd{i}=.S;
end;
```

**Variable Name** M1110384 - M1110304

**Variable Label** Whether had accident in past year that required hospitalization

**Unit of Observation** Individual

**Description** Indicates whether person had accident in past year that required he stay overnight in a clinic or hospital. This information is collected for employed persons only. In 1984-1987, 1990, 1993, 2000-2004 questions related to this topic were not asked.

**Method** Transcribed variable.

**Format**
-2 = N/A - Child / does not apply currently not employed
-1 = Item non-response
0 = Had no accident that required overnight stay in a hospital
1 = Had accident that required overnight stay in a hospital

The original survey variables provided below can be found in the file _P.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984-2004** algorithm:

```
null=.S;
array s03{*}  null        null        null        DP74       EP71       FP87       null       HP82
              IP82        null        KP89        LP95       MP81       NP87       OP76       PP104
              null        null        null        null       null;
array m03{*} M1110384 M1110385 M1110386 M1110387 M1110388 M1110389 M1110390 M1110391
              M1110392 M1110393 M1110394 M1110395 M1110396 M1110397 M1110398 M1110399
              M1110300 M1110301 M1110302 M1110303 M1110304;
do i = 1 to dim(netto);
  if netto{i} in (1,5) then do;
  if s03{i} in (1,2) then m03{i}=1;
  if s03{i} in (3) then m03{i}=0;
  if s03{i} in (-1,-3) then m03{i}=-1;
  if s03{i} in (-2) then m03{i}=-2;
end;
else m03{i}=-2;
end;
```

**Variable Name** M1110484 - M1110404

**Variable Label** Frequency play sports or exercise

**Unit of Observation** Individual

**Description** Indicates how often person plays sports, exercises or walks or swims. In 1987, 1989, 1991, 1993, 2000, 2002 and 2004 questions related to this topic were not asked.

**Method** Transcribed variable.

**Format:**
-2 = N/A - Child
-1 = Item non-response
1 = Play sport or exercise once a year or less, almost never, or never
2 = Play sport or exercise several times a year
3 = Play sport or exercise at least once a month
4 = Play sport or exercise at least once a week

The original survey variable provided below can be found in the _P files.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984-2004** algorithm:

```
*** First code, generate one sport variable for east and west germans ***;
if gsampreg = 2 and zp0203 ge 0 then GP0413 = zp0203+1;
if gsampreg = 2 and zp0203 lt 0 then GP0413 = zp0203;
array sport{*}  AP0202   BP0703   CP0903   null     EP0903   null     GP0413   null
                IP0703   null     KP1203   LP0613   MP0503   NP0303   OP0613   PP0303
                null     RP0303   null     TP1414   null;
array nsport{*} M1110484 M1110485 M1110486 M1110487 M1110488 M1110489 M1110490 M1110491
                M1110492 M1110493 M1110494 M1110495 M1110496 M1110497 M1110498 M1110499
                M1110400 M1110401 M1110402 M1110403 M1110404;
do i = 1 to dim(netto);
 if netto{i} in (1,5) then do;
 if sport{i}=-1 then nsport{i}=-1;
 if sport{i}=-2 then nsport{i}=-2;
 if i = 1 then do;
 if sport{i} gt 0 then nsport{i}=sport{i};
end;

if i in (2,3,5,9,11,13,14,16,18) then do;
 if sport{i} =1 then nsport{i}=4;
 if sport{i} =2 then nsport{i}=3;
 if sport{i} =3 then nsport{i}=2;
 if sport{i} =4 then nsport{i}=1;
end;

if i in (7,12,15,20) then do;
 if sport{i} in (1,2) then nsport{i}=4;
 if sport{i} =3 then nsport{i}=3;
 if sport{i} =4 then nsport{i}=2;
 if sport{i} =5 then nsport{i}=1;
end;
end;
else nsport{i}=-2;
end;
```

**Variable Name** M1111384 - M1111304

**Variable Label** Need help to climb stairs

**Unit of Observation** Individual

**Description** Indicates whether person has trouble with or needs help of others to climb stairs. In 1984-2001 and 2003 questions related to this topic were not asked.

**Method** Transcribed variable.

**Format:**
-2 = N/A - Child
-1 = Item non-response
0 = Doesn't have trouble with stairs or need help with stairs
1 = Has trouble with stairs or needs help of others with stairs

The original survey variable provided below can be found in the _P files.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984-2004** algorithm:

```
null=.-2;
*First set missing values for all other data;
M1111384 - M1111301, M1111303= -2;

array hlpstair{*}        SP87        UP84;
array hlthaffstair {*} M1111302 M1111304;
do i = 1 to dim(hlpstair);
 if hlpstair{i} in (1, 2) then hlthaffstair{i} = 1;
 if hlpstair{i} = 3 then hlthaffstair{i}=0;
 if hlpstair{i} < 0 then hlthaffstair{i}=.-1;
end;
```

**Variable Name** M1111584 - M1111504

**Variable Label** Have difficulty or need help of others to dress

**Unit of Observation** Individual

**Description** Indicates whether person has trouble with or needs help of others to dress.
In 1984-1990 questions related to this topic were not asked.

**Method** Transcribed variable.

**Format:**
-2 = N/A - Child
-1 = Item non-response
0 = Doesn't have trouble with or need help of others to dress
1 = Has trouble with or needs help of others to dress

The variable provided below are derived from the original survey variable "multgrad" which can be found in the file PFLEGE.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984-2004** algorithm:

Note that, in the original survey data the questions in the SOEP asks about differing levels of help required in different years. Please refer to the "PFLEGE.DOC" document. The question from the original survey starts in 1991 and identifies individuals who "require help only with household chores, meals, and drinking."

```
if erhebj=&year;                                                    (&year= 84, 85, …, 04)
if multgrad in (-1,-2,-3) then multgrad=9; char=put(multgrad,5.);
simp&year.=0;
 if substr(char,4,1)=1 then simp&year.=1;
 if multgrad in (9) then do;
 simp&year.=-1;
end;

array simp{*} null     null     null     null     null     null     null     simp91
              simp92   simp93   simp94   simp95   simp96   simp97   simp98   simp99
              simp00   simp01   simp02   simp03   simp04;
array M15{*} M1111584 M1111585 M1111586 M1111587 M1111588 M1111589 M1111590 M1111591
              M1111592 M1111593 M1111594 M1111595 M1111596 M1111597 M1111598 M1111599
              M1111500 M1111501 M1111502 M1111503 M1111504;
do i = 1 to dim(netto);
if netto{i} in (1,5) then do;
 m15{i}=0;
 if simp{i} in (1) then m15{i}=1;
 if simp{i} in (.M) then m15{i}=-1;
end;
else m15{i}= -2;
if i in (1,2,3,4,5,6,7) then m15{i}=.-2;
end;
```

**Variable Name** M1111684 - M1111604

**Variable Label** Have difficulty or need help of others to get in/out of bed

**Unit of Observation** Individual

**Description** Indicates whether person has trouble with or needs help of others to get into/out of bed. In 1984 questions related to this topic were not asked.

**Method** Transcribed variable.

**Format:**
-2 = N/A - Child
-1 = Item non-response
0 = Doesn't have trouble with or need help of others to get in/out of bed
1 = Has trouble with or needs help of others to get in/out of bed

The variables provided below are derived from the original survey variable "multgrad" and "maxgrad" which can be found in the file PFLEGE.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984-2004** algorithm:

```
if erhebj=&year;                                                    (&year= 84, 85, …, 90)
if maxgrad=1 then outh&year.=1;
if maxgrad=3 then do; inh&year.=1; outh&year.=1; end;
if maxgrad=5 then do; bed&year.=1; inh&year. =1; outh&year.=1; end;
if maxgrad in (-1,-3) then do; bed&year.=-1; inh&year.=-1; outh&year.=-1; end;
end;

if erhebj=&year;                                                    (&year= 91, 92, …, 04)
if multgrad in (-1,-2,-3) then multgrad=9; char=put(multgrad,5.);
bed&year.=0;
if substr(char,5,1)=1 then bed&year. =1;
 if multgrad in (9) then do;
  bed&year.=-1;
end;

array bed{*}  null          bed85       - bed04;
array M16{*} M1111684 M1111685 - M1111604;
do i = 1 to dim(netto);
 if netto{i} in (1,5) then do;
 m16{i}=0;
 if bed{i} in (1) then m16{i}=1;
 if bed{i} in (.M) then m16{i}=.M;
end;
else m16{i}=.S;
if i=1 then m16{i}=.S;
end;
```

**Variable Name** M1111784 - M1111704

**Variable Label** Have difficulty or need help of others to shop

**Unit of Observation** Individual

**Description** Indicates whether person has trouble with or needs help of others to shop.
In 1984 questions related to this topic were not asked.

**Method** Transcribed variable.

**Format:**
-2 = N/A - Child
-1 = Item non-response
0 = Doesn't have trouble with or need help of others to shop
1 = Has trouble with or needs help of others to shop

The variables provided below are derived from the original survey variable "multgrad" and "maxgrad" which can be found in the file PFLEGE.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984 2004** algorithm:

```
if erhebj=&year;                                                    (&year= 84, 85, …, 90)
if maxgrad=1 then outh&year.=1;
if maxgrad=3 then do; inh&year.=1; outh&year.=1; end;
if maxgrad=5 then do; bed&year.=1; inh&year. =1; outh&year.=1; end;
if maxgrad in (-1,-3) then do; bed&year.=-1; inh&year.=-1; outh&year.=-1; end;
end;

if erhebj=&year;                                                    (&year= 91, 92, …, 04)
if multgrad in (-1,-2,-3) then multgrad=9; char=put(multgrad,5.);
outh&year.=0;
 if substr(char,1,1)=1 then outh&year.=1;
if multgrad in (9) then do;
 outh&year.=-1;
end;

array outh{*} null          outh85      - outh04;
array M17{*} M1111784 M1111785 - M1111704;
do i = 1 to dim(netto);
if netto{i} in (1,5) then do;
m17{i}=0;
if outh{i} in (1) then m17{i}=1;
if outh{i} in (.M) then m17{i}=.M;
end;
else m17{i}=.S;
if i=1 then m17{i}=.S;
end;
```

**Variable Name** M1111984 - M1111904

**Variable Label** Difficulty doing housework

**Unit of Observation** Individual

**Description** Indicates whether person has trouble doing housework.
In 1984 questions related to this topic were not asked.

**Method** Transcribed variable.

**Format:** -2 = N/A - Child
-1 = Item non-response
0 = Doesn't have trouble doing housework
1 = Has trouble doing housework.

The variables provided below are derived from the original survey variable "mult-grad" and "maxgrad" which can be found in the file PFLEGE.
This algorithm omits individuals with survey non-responses.

**Algorithm**
**1984-2004** algorithm:

```
if erhebj=&year;                                          (&year= 84, 85, …, 90)
if maxgrad=3 then do; inh&year.=1; end;
if maxgrad in (-1,-3) then do; inh&year.=-1;  end;
end;

if erhebj=&year;                                          (&year= 91, 92, …, 04)
if multgrad in (-1,-2,-3) then multgrad=9; char=put(multgrad,5.);
inh&year.=0;
if substr(char,2,1)=1 then inh&year. =1;
if multgrad in (9) then do;
inh&year.=-1;
end;

array inh{*}   null          inh85       – inh04;
array M19{*} M1111984 M1111985 – M1111904;
do i = 1 to dim(netto);
if netto{i} in (1,5) then do;
m19{i}=0;
if inh{i} in (1) then m19{i}=1;
if inh{i} in (.M) then m19{i}=.M;
end;
else m19{i}=.S;
if i=1 then m19{i}=.S;
end;
```

| | |
|---|---|
| **Variable Name** | M1112284 - M1112204 |
| **Variable Label** | Height in centimeters |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates each person's height in centimeters on the interview date.<br>In 1984-2001 and 2003 questions related to this topic were not asked. |
| **Method** | Transcribed variable. |
| **Format** | -2 = N/A - Child<br>-1 = Item non-response<br>0 to 220 = Height in centimeters<br><br>The original survey variables provided below can be found in the file _P.<br>This algorithm omits individuals with |

**Algorithm**

**1984-2001, 2003** algorithm: Data not available in SOEP

**2002** algorithm: M1112202=sp90
**2004** algorithm: M1112204=up90

| | |
|---|---|
| **Variable Name** | M1112384 - M1112304 |
| **Variable Label** | Weight in kilos |
| **Unit of Observation** | Individual |
| **Description** | This variable indicates each person's weight in kilos on the interview date.<br>In 1984-2001 and 2003 questions related to this topic were not asked. |
| **Method** | Transcribed variable. |
| **Format** | -2 = N/A - Child<br>-1 = Item non-response<br>0 to 400 = Weight in kilos<br><br>The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses. |

**Algorithm**

**1984-2001, 2003** algorithm: Data not available in SOEP

**2002** algorithm: M1112302=sp91
**2004** algorithm: M1112304=up91

**Variable Name** M1112484 - M1112404

**Variable Label** Disability Status of Individual

**Unit of Observation** Individual

**Description** This variable indicates disability status at the time of the survey for all individuals in the household 16 years of age and older.

**Method** Disability is here defined as a share of legally attested disability of more than 30%. In 1986, 1990 and 1993 this information was in SOEP not asked. If in the following year a legally attested disability existed then this information is assigned to the previous year.

**Format**
-2 = N/A – Child
-1 = Item non-response
0 = Not disabled
1 = Disabled

The original survey variables provided below can be found in the file _P.
This algorithm omits individuals with survey non-responses.

**Algorithm**

Degree of disability variable list by survey year - each entry denoted in algorithm as Ydisab:
ap5202 bp7402 cp74b02 dp0602 ep6602 fp8202 gp7702 hp7702 ip7802 jp7802 kp8402 lp9102 mp7702 np8102 op7002 pp9702 qp9702 rp9702 sp9502 tp9902 up9202

**1984-2004** algorithm:

*This step is to assign a value for 1990 using 1991 degree of disability;
if gnetto in (1,2,3,4) and ghnetto in (1) then do; gp7702 = 0; if hp7702 > 0 then gp7702=hp7702; end;

*This step is to assign a value for 1993 using 1992 degree of disability;
if jnetto in (1,2,3,4) and jhnetto in (1) then do; jp7802 = 0; if ip7802 > 0 then jp7802=ip7802; end;

*This step is to assign values to respondents in 1986 missing degree of disability information;
if cnetto in (1,2,3,4) and chnetto in (1) then do; if cp74b02<0 then do; if bp7402 > 0 then cp74b02 = bp7402; end;

*Next assign disability status for all individuals using the variable list shown above;

if Ydisab ▲◄ 30 then D11110$$=1; (Y=a-u, $$=84-04)
else D11110$$=0;

**Variable Name** M1112584 - M1112504

**Variable Label** Satisfaction With Health

**Unit of Observation** Individual

**Description** This variable indicates satisfaction with health at the time of the survey for all individuals in the household 16 years of age and older.

**Method** Individuals are asked to report how satisfied they are with their health on a scale of 0 to 10, in which 0 means totally unhappy and 10 means totally happy. Individuals reporting 0, 1, 2 or 3 are considered to be not satisfied with their health. Individuals reporting 4, 5, 6, 7, 8, 9, or 10 are considered to be satisfied with their health.

**Format**

-2 = N/A - Child
-1 = Item non-response
0 = Not Satisfied
1 = Satisfied

The original survey variables provided below can be found in the file __P.
This algorithm omits individuals with survey non-responses.

Equivalent Data File Variable Definitions: D11101__ = Age of Individual

**Algorithm**

Satisfaction With Health variable list by survey year - each entry denoted in algorithm as Ysatis

ap0301 bp0101 cp0101 dp0101 ep0101 fp0101 gp0101 hp1001 ip9801 jp0101 kp0101 lp0101 mp0101 np0101 op0101 pp0101 qp0101 rp0101 sp0101 tp0101 up0101;

```
if D11101$$ ge 16 then do                    (Y=a-u, $$=84-04)
                    if Ysatis = 0,1,2, or 3 then D11111$$ = 0
                    else if Ysatis = 4, 5, 6, 7, 8, 9, or 10 then D11111$$ = 1
                    else D11111$$ = .M
                    end
                    if D11101$$ lt 16 then D11111$$ = .C
```

**Variable Name** M1112684 - M1112604

**Variable Label** Self-Rated Health Status

**Unit of Observation** Individual

**Description** This variable indicates each person's self-rated health status.

**Method** Transcribed variable.

**Format:**
-2 = N/A - Child
-1 = Item non-response
1 = Excellent
2 = Very good
3 = Good
4 = Fair
5 = Poor

The original survey variables provided below can be found in the file _P.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-1991** algorithm: Data not available in SOEP
**1992-2004**

```
array sta1{*} null    null    null    null    null    null    null    null
              ip77    null    kp83    lp89    mp75    np79    op66    pp95
              qp95    rp95    sp86    tp98    up83;
array stat{*} M1112684 M1112685 M1112686 M1112687 M1112688 M1112689 M1112690 M1112691
              M1112692 M1112693 M1112694 M1112695 M1112696 M1112697 M1112698 M1112699
              M1112600 M1112601 M1112602 M1112603 M1112604;
do i = 1 to dim(netto);
if netto{i} in (1,5) then do;
if sta1{i} ge 0 then stat{i}=sta1{i};
if sta1{i} in (-1,-3) then stat{i}=-1;
if sta1{i} in (-2) then stat{i}=-2;
end;
else stat{i}=-2;
end;
```

**Variable Name** M1112784 - M1112704

**Variable Label** Number of doctor visits in previous year

**Unit of Observation** Individual

**Description** Number of doctor visits in previous year. The SOEP asked for the number of trips to the doctor's in the last three months. This information is multiplied by 4 to get the annual figure. In the years 1984 to 1987 and 1994 the SOEP asked for various medical specialist. This information was add up to yield one measure.
In 1990 and 1993 the Number of doctor visits were not asked.

**Method** Transcribed variable

**Format:** -2 = N/A - Child
-1 = Item non-response
0 to 400 = doctors visits in previous year.

The original survey variables provided below can be found in the file _P.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1990 and 1993** algorithm: Data not available in SOEP
**1984-1989, 1991-1992, 1994-2004** algorithm:

```
ap5001 = sum(ap5002-ap5012);
bp7101 = sum(bp7102-bp7112);
cp7101 = sum(cp7102-cp7112);
dp7101 = sum(dp7102-dp7112);
kp8601 = sum(kp8602-kp8612);

array s100{*} ap5001     bp7101     cp7101     dp7101     EP6801     FP8401     null       HP7901
              IP7901     null       kp8601     LP9201     MP7801     NP8201     OP7101     PP9801
              QP9801     RP9801     SP9601     TP10001    up9301;
array s101{*} M1112784 – M1112704;
do i = 1 to dim(netto);
 if netto{i} in (1,5) then do;
  if s100{i} ge 0       then s101{i}=4*s100{i};
  if s100{i} in (-1,-3) then s101{i}=.M;
  if s100{i} in (-2)    then s101{i}=0;
 end;
 else s101{i}=.S;
end;
```

| | | |
|---|---|---|
| **Variable Name** | M1110584 - M1110504 | Variable Label: Have had stroke |
| | M1110684 - M1110604 | Variable Label: High blood pressure/circulation problems |
| | M1110784 - M1110704 | Variable Label: Have or had diabetes |
| | M1110884 - M1110804 | Variable Label: Have or had cancer |
| | M1110984 - M1110904 | Variable Label: Have or had psychiatric problems |
| | M1111084 - M1111004 | Variable Label: Have or had arthritis |
| | M1111184 - M1111104 | Variable Label: Angina or heart condition |
| | M1111284 - M1111204 | Variable Label: Have or had asthma or breathing difficulty |
| | M1111484 - M1111404 | Variable Label: Have difficulty or need help to bathe |
| | M1111884 - M1111804 | Variable Label: Walk 10+ minutes difficult |
| | M1112084 - M1112004 | Variable Label: Health limits bending, lifting, stooping |
| | M1112184 - M1112104 | Variable Label: Health limits vigorous physical activity |

**Algorithm**
**1984-2004** algorithm: Data not available in SOEP

**Variable Name** W1110184 - W1110104

**Variable Label** Individual Weight

**Unit of Observation** Individual

**Description** Indicates the individual's sample weight.

**Method** Individual weights to compensate for unequal probabilities of selection and sample attrition are necessary to obtain populations based statistics.

**Format** The value of this variable ranges from 0 to 80.000.

The original survey variables provided below can be found in the file PHRF.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: W11101$$ = Yphrf ($$=84-04, Y=a-u)

**Variable Name** W1110284 - W1110204

**Variable Label** Household Weight

**Unit of Observation** Household

**Description** Indicates the household's sample weight.

**Method** Household weights to compensate for unequal probabilities of selection and sample attrition are necessary to obtain populations based statistics.

**Format** The value of this variable ranges from 0 to 80.000.

The original survey variables provided below can be found in the file HHRF.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984-2004** algorithm: W11102$$ = Yhhrf ($$=84-04, Y=a-u)

**Variable Name** W1110384 - W1110304

**Variable Label** Longitudinal Weight

**Unit of Observation** Individual

**Description** Indicates the individual's longitudinal sample weight.

**Method** Longitudinal weights in the SOEP are constructed using inverse staying factors. The staying factor is the inverse of the probability that an individual participated in the named year. By themselves, W11103$$ are only staying factors.

To properly weight a balanced sample of individuals represented in five years of the survey (1986 through 1990), for example, create a longitudinal weight by multiplying the individual weight in 1986 by the staying factors in 1987, 1988, 1989, and 1990.

Example:
longitudinal weight = W1110186*W1110387*W1110388*W1110389*W1110390

The created longitudinal weight should be used with any longitudinal sample.

**Format** N/A

The original survey variables provided below can be found in the file PHRF.
This algorithm omits individuals with survey non-responses.

**Algorithm**

**1984** algorithm: N/A
**1985-2004** algorithm: W11103$$ = Ypbleib ($$=85-04, Y=b-u)

| | |
|---|---|
| **Variable Name** | W1110484 - W1110404 |
| **Variable Label** | Population Factor |
| **Unit of Observation** | Household |
| **Description** | This correction factor is already integrated into W11101$$ and is not repeated here. |
| **Format** | -1 = no information available |

| | |
|---|---|
| **Variable Name** | W1110584 - W1110504 |
| **Variable Label** | Individual Weight - Immigrant Sample |
| **Unit of Observation** | Individual |
| **Description** | Indicates the individual's weight for the immigrant sample. |
| **Method** | Individual weights to compensate for unequal probabilities of selection and sample attrition are necessary to obtain populations based statistics.<br><br>These weights should be used when analyses include only the immigrant sample (D).<br><br>For a more detailed discussion of weighting issues in the SOEP please see *Desktop Companion to the German Socio-Economic Panel Study (SOEP*). |
| **Format** | The value of this variable ranges from 0 to 80.000.<br><br>The original survey variables provided below can be found in the file PHRF.<br>This algorithm omits individuals with survey non-responses. |
| **Algorithm** | |
| **1984-1994** | N/A |
| **1995-2004 algorithm**: | W11105$$ = Yphrfd ($$=95-04, Y=l-u) |

| | |
|---|---|
| **Variable Name** | W1110684 - W1110604 |
| **Variable Label** | Household Weight - Immigrant Sample |
| **Unit of Observation** | Household |
| **Description** | Indicates the household's weight for the immigrant sample. |
| **Method** | Household weights to compensate for unequal probabilities of selection and sample attrition are necessary to obtain populations based statistics.<br><br>These weights should be used when analyses include only the immigrant sample (D).<br><br>For a more detailed discussion of weighting issues in the SOEP please see *Desktop Companion to the German Socio-Economic Panel Study (SOEP*). |
| **Format** | The value of this variable ranges from 0 to 80.000.<br><br>The original survey variables provided below can be found in the file HHRF.<br>This algorithm omits individuals with survey non-responses. |
| **Algorithm** | |
| **1984-1994** | N/A |
| **1995-2004** algorithm: | W11105$$ = Yhhrfd*(1.05)  ($$=95-04, Y=l-u) |

| Variable Name | Variable Label: |
|---|---|
| W1110784 - W1110704 | Cross-sectional Weight - Enumerated Individuals |
| W1110884 - W1110804 | Longitudinal Weight - Enumerated Individuals |
| W1110984 - W1110904 | Population Factor for W11103$$ |
| W1111084 - W1111004 | Population Factor for W11107$$ |
| W1111184 - W1111104 | Population Factor for W11108$$ |

**Algorithm**

**1984-2004** algorithm: Data not available in SOEP

## References


*Aldi, J.M., Hagenaars, Klaus de Vos, M. Asghar Zaidi (1994). "Poverty Statistics in the late 1980's: Research Based on Micro-Data." Study carried out for Eurostat. Luxemborg, Office for Official Publications of the European Communities. Population and Social Conditions Series 3C, p. 18.*

*Betson, David, and Robert Michael. (1993). "A Recommendation for the Construction of Equivalence Scales." Unpublished memorandum prepared for the Panel on Poverty and Family Assistance, Committee on National Statistics, National Research Council. Department of Economics, University of Notre Dame.*

*Frick, Joachim R. und Grabka, Markus M. (2001): Der Einfluß von Imputed Rent auf die personelle Einkommensverteilung. In: Jahrbücher für Nationalökonomie und Statistik, 221(3), S. 285-308.*

*Frick, Joachim R. und Grabka, Markus M. (2003): Imputed Rent and Income Inequality: a Decomposition Analysis for the UK, West Germany and the USA. Review of Income and Wealth Vol. 49(4), p. 513-537.*

*Frick, Joachim R. and Markus M. Grabka (2003): Imputed Rent and Income Inequality: a Decomposition Analysis for the UK, West Germany and the USA. Review of Income and Wealth Vol 49(4), p. 513-537.*

*Grabka, M.M. and Frick, J.R. (2003): Imputation of Item-Non-Response on Income Questions in the SOEP 1984–2002. DIW Research Note No. 29, Berlin October 2003. (http://www.diw.de/deutsch/produkte/publikationen/materialien/docs/papers/diw_rn03-10-29.pdf).*

*Little, Roderick J.A. and Hong-Lin Su. (1989). "Item Non-response in Panel Surveys." in Panel Surveys. (D. Kasprzyk, G. Duncan, and M.P. Singh, eds.). New York: John Wiley.*

*Oaxaca, R. (1973): Sex Discrimination in Wages. In: Ashenfelter, O. and Rees, A. (Eds.): Discrimination in Labor Markets, Princeton, NJ: Princeton University Press, 124-151.*

*Organization for Economic Cooperation and Development. (1982). The OECD List of Social Indicators. Paris: Organization for Economic Cooperation and Development.*

*Schwarze, Johannes (1995): Simulating German Income and Social Security Tax Payments Using The SOEP. Cross-National Studie in Aging. Programm Project Paper No. 19. All-University Gerontology Center Maxwell School of Citizenship and Public Affairs, Syracuse University, New York.*

*United Nations (1977): Provisional Guide-lines on Statistics of the Distribution of Income, Consumption and Accumulation of Households. Studies in Methods. Series M. No. 61. New York.*

*Yates, Judith (1994): Imputed Rent and Income Distribution. Review of Income and Wealth, 40 (1): 43-66.*