

Grabka, Markus M.

Research Report

Codebook for the \$PEQUIV File 1984-2009: CNEF Variables with Extended Income Information for the SOEP

DIW Data Documentation, No. 49

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Grabka, Markus M. (2010) : Codebook for the \$PEQUIV File 1984-2009: CNEF Variables with Extended Income Information for the SOEP, DIW Data Documentation, No. 49, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/129254>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

Data Documentation

49

Markus M. Grabka

**Codebook for the
\$PEQUIV File 1984-2009
CNEF Variables with Extended
Income Information for the SOEP**

Berlin, June 2010

IMPRESSUM

© DIW Berlin, 2010

DIW Berlin
Deutsches Institut für Wirtschaftsforschung
Mohrenstr. 58
10117 Berlin
Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
www.diw.de

ISSN 1861-1532

All rights reserved.
Reproduction and distribution
in any form, also in parts,
requires the express written
permission of DIW Berlin.

Data Documentation 49

Markus M. Grabka

Codebook for the \$PEQUIV File 1984-2009

CNEF variables with extended income information for the SOEP

Berlin, June 2010

* DIW Berlin, Längsschnittstudie "Sozio-oekonomisches Panel" (SOEP).
Homepage: <http://www.diw.de/soep/>

Kontakt: mgrabka@diw.de; phone: +49-30-897-89-339.

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, Statistics Canada, the Melbourne Institute of Applied Economics and Social Research (MI), the Korea Labor Institute and the Swiss Foundation for research in Social Sciences (FORS) funded by the National Institute on Aging and by the DIW Berlin. For extensive documentation of the CNEF cf.

<http://www.human.cornell.edu/che/PAM/Research/Centers-Programs/German-Panel/cnef.cfm>

or:

Joachim R. Frick, Stephen P. Jenkins, Dean R. Lillard, Oliver Lipps, and Mark Wooden (2007):
The Cross-National Equivalent File (CNEF) and its Member Country Household Panel Studies.
In: Schmollers Jahrbuch (Journal of Applied Social Science Studies), Vol. 127(4) , p. 627-654.

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 2010 release of the \$PEQUIV-files has been updated to include the 2009 (wave Z) SOEP data.
- Population for \$PEQUIV is made up by all members of households who were successfully interviewed (i.e., persons with \$NETTO-codes 10 to 39 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 and onwards (i.e., variables w11105\$\$ and w11102\$).

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: 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.
- Any missing income information due to partial unit non response (PUNR, non responding individuals in households with at least one successful interview) has been imputed according to the longitudinal and cross-sectional imputation procedures described in: Frick, J.R.; M.M. Grabka and O. Groh-Samberg (2010): Dealing with Incomplete Household Panel Data in Inequality Research. SOEP Papers on Multidisciplinary Panel Data Research at DIW Berlin, No. 290, Berlin (see: http://www.diw.de/documents/publikationen/73/diw_01.c.354683.de/diw_sp0290.pdf). Due to lacking detailed information about income receipt, only six income components have been imputed: individual labour income (I11110\$\$), social security pensions (I11108\$\$), unemployment benefits (IUNBY\$\$), maternity benefits (IMATY\$\$), student grants (ISTUY\$\$) and private transfers (IELSE\$\$). This information is also used to generate a more thorough measure for taxes and social contributions paid by private households.
- 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.

General variable naming conventions for the \$PEQUIV-variables: (see variable list on page 4):

- Variable names are longitudinally consistent using a two-digit suffix – instead of a four-digit suffix used in the original CNEF-files – indicating the survey year (wave A = 84, wave B = 85, ..., wave Z = 09, \$\$ = 84, 85, ..., 09).

Variable naming conventions for the single income components:

- 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 1984-2009

Label	Variable List	Page
Identifiers:		
Unique Person Number	X11101LL	10
Household Identification Number	X11102\$\$	11
Individual in Household at Survey	X11103\$\$	12
Sub-sample Identifier	X11104LL	13
Person in Household Interviewed	X11105\$\$	14
Demographics :		
Age of Individual	D11101\$\$	15
Sex of Individual	D11102LL	16
Race of Household Head	D11103\$\$	17
Marital Status of Individual	D11104\$\$	18
Relationship to Household Head	D11105\$\$	19
Number of Persons in Household	D11106\$\$	20
Number of Children in Household	D11107\$\$	21
Education With Respect to High School	D11108\$\$	22
Number of Years of Education	D11109\$\$	23
Race of Individual	D11112LL	24
Employment:		
Annual Work Hours of Individual	E11101\$\$	25
Impute Annual Work Hours of Individual	E11201\$\$	26
Employment Status of Individual	E11102\$\$	27
Employment Level of Individual	E11103\$\$	28
Primary Activity of Individual	E11104\$\$	29
Occupation of Individual	E11105\$\$	30
1 Digit Industry Code of Individual	E11106\$\$	31
2 Digit Industry Code of Individual	E11107\$\$	32
Equivalence scale inputs:		
Number HH members age 0-14	H11101\$\$	34
Number HH members age 15-18	H11102\$\$	34
Number HH members age 0-1	H11103\$\$	34
Number HH members age 2-4	H11104\$\$	34
Number HH members age 5-7	H11105\$\$	34
Number HH members age 8-10	H11106\$\$	34
Number HH members age 11-12	H11107\$\$	34
Number HH members age 13-15	H11108\$\$	34
Number HH members age 16-18	H11109\$\$	34
Number HH members age 19+ or 16-18 and indep.	H11110\$\$	34
Indicator - Wife/spouse in Household	H11112\$\$	36
Equivalence Weight algorithms:		
OECD Equivalence Weight		37
Modified OECD Equivalence Weight		37
Other Equivalence Weights		37
Location:		
Area of Residence	L11101\$\$	38
Region of Residence	L11102\$\$	39

Label	Variable List	Page
Macro-level Variables:		
Consumer Price Index	Y11101\$\$	40
Aggregated Income variables:		
Household Pre-Government Income	I11101\$\$	41
Household Post-Government Income	I11102\$\$	42
Household Labor Income	I11103\$\$	43
Household Asset Income	I11104\$\$	44
Household Imputed Rental Value	I11105\$\$	45
Household Private Transfers	I11106\$\$	46
Household Public Transfers	I11107\$\$	47
Household Social Security Pensions	I11108\$\$	48
Total Household Taxes	I11109\$\$	49
Individual Labor Earnings	I11110\$\$	50
Household Federal Taxes	I11111\$\$	51
Household Social Security Taxes	I11112\$\$	52
Household Post-Government Income (TAXSIM)	I11113\$\$	53
Total Household Taxes (TAXSIM)	I11114\$\$	53
Household State Taxes (TAXSIM)	I11115\$\$	53
Household Federal Taxes (TAXSIM)	I11116\$\$	53
Household Private Retirement Income	I11117\$\$	54
Household Windfall Income	I11118\$\$	55
Imputation flag: Household Pre-Government Income	I11201\$\$	56
Imputation flag: Household Post-Government Income	I11202\$\$	57
Imputation flag: Household Labor	I11203\$\$	58
Imputation flag: Household Asset Income	I11204\$\$	59
Imputation flag: Household Imputed Rental Value	I11205\$\$	60
Imputation flag: Household Private Transfers	I11206\$\$	61
Imputation flag: Household Public Transfers	I11207\$\$	62
Imputation flag: Household Social Security Pensions	I11208\$\$	63
Imputation flag: Total Household Taxes	I11209\$\$	64
Imputation flag: Individual Labor Earnings	I11210\$\$	65
Imputation flag: Private Retirement Income	I11217\$\$	66
Imputation flag: Household Windfall Income	I11218\$\$	67
Single income components at the household level:		
Income from rental and leasing	RENTY\$\$	68
Operation, maintenance costs	OPERY\$\$	69
Interest, dividend income	DIVDY\$\$	70
Child allowance	CHSPT\$\$	71
Housing benefit	HOUSE\$\$	72
Nursing allowances	NURSH\$\$	73
Social assistance	SUBST\$\$	74
Social assistance for special circumstances	SPHLP\$\$	75
Social assistance for elderly	SSOLD\$\$	76
Unemployment benefit II	ALG2\$\$	77
Housing support for owner-occupiers	HSUP\$\$	78
Losses from renting and leasing	LOSSR\$\$	79
Losses from capital investment	LOSSC\$\$	80
Additional child benefit	ADCHB\$\$	81

Label	Variable List	Page
Imputation flag: Income from rental and leasing	FRENTY\$\$	82
Imputation flag: Operation, maintenance costs	FOPERY\$\$	83
Imputation flag: Interest, dividend income	FDIVDY\$\$	84
Imputation flag: Child allowance	FCHSPT\$\$	85
Imputation flag: Housing benefit	FHOUSE\$\$	86
Imputation flag: Nursing allowances	FNURSH\$\$	87
Imputation flag: Social assistance	FSUBST\$\$	88
Imputation flag: Social assistance for spec. circumst.	FSPHLP\$\$	89
Imputation flag: Social assistance for elderly	FSSOLD\$\$	90
Imputation flag: Unemployment benefit II	FALG2\$\$	91
Imputation flag: Housing support for owner-occupiers	FHSUP\$\$	92
Imputation flag: Losses from renting and leasing	FLOSSR\$\$	93
Imputation flag: Losses from capital investment	FLOSSC\$\$	94
Imputation flag: Additional child benefit	FADCHB\$\$	95
Single income components at the individual level:		
Wages, Salary from main job	IJOB1\$\$	96
Income from secondary employment	IJOB2\$\$	97
Income from self-employment	ISELF\$\$	98
Combined old-age, disability, etc. pensions	IOLDY\$\$	99
Combined widows and orphans pension	IWIDY\$\$	100
Combined company pension	ICOMP\$\$	101
Combined private pension	IPRVP\$\$	102
Unemployment benefit	IUNBY\$\$	103
Unemployment assistance	IUNAY\$\$	104
Subsistence allowance	ISUBY\$\$	105
Old-age transition benefit	IERET\$\$	106
Maternity benefit	IMATY\$\$	107
Student grants	ISTUY\$\$	108
Military community service pay	IMILT\$\$	109
Alimony	IALIM\$\$	110
Private Transfers received	IELSE\$\$	111
13th monthly salary	I13LY\$\$	112
14th monthly salary	I14LY\$\$	113
Christmas bonus	IXMAS\$\$	114
Vacation bonus	IHOLY\$\$	115
Profit-sharing	IGRAY\$\$	116
Other bonuses	IOTHY\$\$	117
Commuting expenses, travel grant	ITRAY\$\$	118
Indemnity payments	IDEMY\$\$	119
Statutory pension insurance	IGRV1\$\$	120
Social miners insurance pension	ISMP1\$\$	121
Civil servant pension	ICIV1\$\$	122
War victim pension	IWAR1\$\$	123
Farmer Pension	IAGR1\$\$	123
Statutory accident insurance pension	IGUV1\$\$	124
Supplementary benefits for civil servants	IVBL1\$\$	125
Company pension	ICOM1\$\$	126
Private pension	IPRV1\$\$	127
Other pension	ISON1\$\$	128
Widows and orphans stat. pension insurance	IGRV2\$\$	130
Widows / orphans social miners insurance pens.	ISMP2\$\$	131

Label	Variable List	Page
Widows / orphans civil servant pension	ICIV2\$\$	132
Widows / orphans war victim pension	IWAR2\$\$	133
Widows / orphans farmer pension	IAGR2\$\$	134
Widows / orphans statutory accident insurance	IGUV2\$\$	135
Widows / orphans supplement. benefits	IVBL2\$\$	136
Widows / orphans company pension	ICOM2\$\$	137
Widows / orphans private pension	IPRV2\$\$	138
Other / orphans widows pension	ISON2\$\$	139
Imputation flag: Wages, Salary from main job	FJOB1\$\$	140
Imputation flag: Income from secondary job	FJOB2\$\$	141
Imputation flag: Income from self-employment	FSELF\$\$	142
Imputation flag: Unemployment benefit	FUNBY\$\$	143
Imputation flag: Old-age, civil servants pens.	FOLDY\$\$	144
Imputation flag: Widows / orphans pension	FWIDY\$\$	145
Imputation flag: Combined company pension	FCOMP\$\$	146
Imputation flag: Combined private pension	FPRVP\$\$	147
Imputation flag: Unemployment assistance	FUNAY\$\$	148
Imputation flag: Subsistence allowance	FSUBY\$\$	149
Imputation flag: Old-age transition benefit	FERET\$\$	150
Imputation flag: Maternity benefit	FMATY\$\$	151
Imputation flag: Student grants	FSTUY\$\$	152
Imputation flag: Military community service pay	FMILT\$\$	153
Imputation flag: Alimony	FALIM\$\$	154
Imputation flag: Private Transfers received	FELSE\$\$	155
Imputation flag: 13th monthly salary	F13LY\$\$	156
Imputation flag: 14th monthly salary	F14LY\$\$	157
Imputation flag: Christmas bonus	FXMAS\$\$	158
Imputation flag: Vacation bonus	FHOLY\$\$	159
Imputation flag: Profit-sharing	FGRAY\$\$	160
Imputation flag: Other bonuses	FOTHY\$\$	161
Imputation flag: Commuting, travel grant	FTRAY\$\$	162
Imputation flag: Indemnity	FDEMY\$\$	163
Imputation flag: Statutory pension insurance	FGRV1\$\$	164
Imputation flag: Widows stat. pension insurance	FGRV2\$\$	165
Medical/health:		
Whether spent night in hospital in last year	M11101\$\$	166
Number of nights (days) spent in hospital	M11102\$\$	167
Whether had accident that required hospital	M11103\$\$	168
Frequency of sports or exercise	M11104\$\$	169
Have had stroke	M11105\$\$	170
Have or had high blood pressure/hypertension	M11106\$\$	171
Have or had diabetes	M11107\$\$	172
Have or had cancer	M11108\$\$	173
Have or had psychiatric problems	M11109\$\$	174
Have or had arthritis	M11110\$\$	175
Have or had angina or heart condition	M11111\$\$	176
Have or had asthma or breathing difficulties	M11112\$\$	177
Have trouble climbing stairs	M11113\$\$	178
Have trouble with bath	M11114\$\$	179
Have trouble dressing	M11115\$\$	180
Have trouble getting out of bed	M11116\$\$	181
Have trouble shopping	M11117\$\$	182

Label	Variable List	Page
Have trouble walking	M11118\$\$	183
Have trouble doing housework	M11119\$\$	184
Have trouble bending, lifting, stooping	M11120\$\$	185
Health limits vigorous physical activities	M11121\$\$	186
Height (in meters)	M11122\$\$	187
Weight (in kilos)	M11123\$\$	188
Disability Status of Individual	M11124\$\$	189
Subjective Satisfaction with Health	M11125\$\$	190
Self-Rated Health Status	M11126\$\$	191
Number of doctor visits in previous year	M11127\$\$	192
Psychological information:		
Overall life satisfaction	P11101\$\$	193
Weights:		
Cross-sectional Weight – Respondent Individuals	W11101\$\$	194
Household Weight	W11102\$\$	195
Longitudinal Weight – Respondent Individuals	W11103\$\$	196
Population Factor for W11101\$\$	W11104\$\$	197
Individual Weight – all samples	W11105\$\$	198
Cross-sectional Weight – Enumerated Individuals	W11107\$\$	199
Longitudinal Weight – Enumerated Individuals	W11108\$\$	199
Population Factor for W11103\$\$	W11109\$\$	199
Population Factor for W11107\$\$	W11110\$\$	199
Population Factor for W11108\$\$	W11111\$\$	199
References:		200

Using the \$PEQUIV File Codebook

Variable Name	I11110\$\$	<i>Name of Variable in the \$PEQUIV-File</i>
Variable Label	Labor Earnings of Individual	
Unit of Observation	Individual / Household / Year	
Period	Annual	<i>Periodicity</i>
		<i>Description of Variable Content</i>
Description	This variable represents the labor earnings of each individual in the household.	
Method		<i>Description of Variable Creation</i>
	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		<i>Variable Format in the \$PEQUIV-File</i>
	Not formatted. This variable is in current year EURO.	
Algorithm	$I11110\$ = IJOB1\$ + IJOB2\$ + ISELF\$ + IMILT\$ + I13LY\$ + I14LY\$ + IXMAS\$ + IHOLY\$ + IGRAY\$ + IOTHY\$$	

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	X11101LL = PERSNR

Variable Name	X11102\$\$
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	X11102\$\$ = Yhhnr (Y = a, b , ...)

Variable Name	X11103\$\$
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	If Ynetto >= 10 & Ynetto < 40 and Yhnetto=1 then X11103\$\$ = 1 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	<p>The SOEP contains seven different samples.</p> <ol style="list-style-type: none"> 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 German 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. 8) Sample H is an additional supplemental random sample that represents the population of Germany in 2006. 9) Sample I is an additional supplemental random sample that represents the population of Germany in 2009.
Format	<p>21 = Sample A 22 = Sample B 23 = Sample C 24 = Sample D 25 = Sample E 26 = Sample F 27 = Sample G 28 = Sample H 29 = Sample I</p> <p>The original survey variables provided below can be found in the file PPFAD.</p>
Algorithm	<pre> 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 else if psample = 8 then X11104LL = 28 else if psample = 9 then X11104LL = 29 </pre>

Variable Name	X11105\$\$
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	If Ynetto >= 10 & Ynetto < 30 then X11105\$\$ = 1; else X11105\$\$ = 0;

Variable Name	D11101\$\$
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	<p>-1 = Item non-response 0 = Newborn up to first birthday</p> <p>The value of this variable ranges from 0 to 105.</p> <p>The original survey variables provided below can be found in the file PPFAD. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<p>D11101\$\$ = 19\$\$ - GEBJAHR (\$\$=84-99)</p> <p>D11101\$\$ = 20\$\$ - GEBJAHR (\$\$=00-09)</p>

Variable Name	D11102LL
Variable Label	Gender of Individual
Unit of Observation	Individual
Description	Indicates the gender of the individual.
Method	<p>The SOEP records the gender (SEX) of each individual. This information is acquired once and is not obtained in subsequent years.</p> <p>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.</p>
Format	<p>-1 = Item non-response 1 = Male 2 = Female</p> <p>The original survey variables provided below can be found in the file PPFAD. This algorithm omits individuals with survey non-responses.</p>
Algorithm	if SEX ne -1 then D11102LL = SEX

Variable Name	D11103\$\$
Variable Label	Race of Household Head
Unit of Observation	Individual
Description	Indicates the race of the interviewed head of household.
Method	<p>Race is <u>not</u> available in the SOEP.</p> <p>However, to separate Germans from non-Germans use the variables about</p> <ul style="list-style-type: none"> ○ nationality (NATION\$\$) which can be found in the \$PGEN-files or ○ the information about whether a person was born in Germany (GERMBORN) or ○ the country of origin (CORIGIN) whereas both can be found in the PPFAD-file.
Format	-1 = no information available

Variable Name	D11104\$\$
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	<p>-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)</p> <p>The original survey variables provided below can be found in the file _PGEN. This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions: D11101__ = Age of Individual</p>
Algorithm	<pre> if D11101\$\$ ge 16 then do 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 </pre>

Variable Name	D11105\$\$
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	<p>-1 = Item non-response 1 = Head 2 = Partner 3 = Child 4 = Relative 5 = Non-relative</p> <p>The original survey variables provided below can be found in the file _PBRUTTO. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<pre> if Ystell = 0 then D11105\$\$ = 1 else if Ystell = 1,2,13 then D11105\$\$ = 2 else if Ystell = 3 or 4 then D11105\$\$ = 3 else if Ystell = 5, 6, 7, 8, 9, 10 then D11105\$\$ = 4 else if Ystell =11,12,13 then D11105\$\$ = 5 else D11104\$\$ = -1 </pre>

Variable Name	D11106\$\$
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	D11106\$\$ = Yhhgr

Variable Name	D11107\$\$
Variable Label	Number of Children in Household
Unit of Observation	Household
Description	Indicates the number of persons in the household under age of 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	<p>-1 = Item non-response</p> <p>The value of this variable ranges from 0 to 10.</p> <p>The original survey variables provided below can be found in the file \$PPFAD. This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions: D11101__ = Age of Individual</p>
Algorithm	<pre> if age\$\$ ge 0 and age\$\$ le 17 then sumkids\$\$=1 if age\$\$ = . and \$netto in (20, 21, 22, 23) then sumkids\$\$=1 D11107\$\$ = sum of (sumkids\$\$) in the household </pre>

Variable Name	D11108\$\$						
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	<p>This variable is coded as follows:</p> <table> <tr> <td>Less than = High School</td><td> Intermediate secondary school (Realschule) Lower secondary school (Hauptschule) Other None </td></tr> <tr> <td>High School =</td><td> Upper secondary school degree giving access to university studies (Abitur) Certificate of aptitude for specialized short-course higher education (Fachhochschulreife) Apprenticeship (Lehre) Specialized vocational school (Berufsfachschule) </td></tr> <tr> <td>More than = High School</td><td> School of health care (Schule des Gesundheitswesens) Specialized college of higher education, post-secondary technical (Fachhochschule) College Technical university usually requiring practical training as part of the studies (Technische Universität) Civil service training </td></tr> </table>	Less than = High School	Intermediate secondary school (Realschule) Lower secondary school (Hauptschule) Other None	High School =	Upper secondary school degree giving access to university studies (Abitur) Certificate of aptitude for specialized short-course higher education (Fachhochschulreife) Apprenticeship (Lehre) Specialized vocational school (Berufsfachschule)	More than = High School	School of health care (Schule des Gesundheitswesens) Specialized college of higher education, post-secondary technical (Fachhochschule) College Technical university usually requiring practical training as part of the studies (Technische Universität) Civil service training
Less than = High School	Intermediate secondary school (Realschule) Lower secondary school (Hauptschule) Other None						
High School =	Upper secondary school degree giving access to university studies (Abitur) Certificate of aptitude for specialized short-course higher education (Fachhochschulreife) Apprenticeship (Lehre) Specialized vocational school (Berufsfachschule)						
More than = High School	School of health care (Schule des Gesundheitswesens) Specialized college of higher education, post-secondary technical (Fachhochschule) College Technical university usually requiring practical training as part of the studies (Technische Universität) Civil service training						
Format	<p>-1 = N/A – Child / Item non-response 1 = Less than High School 2 = High School 3 = More than High School</p> <p>The original survey variables provided below can be found in the file _PGEN. This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions: D11101__ = Age of Individual</p>						
Algorithm	<pre> if Ypsbil=.B then Ypsbil=0; 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; </pre>						

Variable Name	D11109\$\$
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	if Ybilzeit=.B then D11109\$\$=0; 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 <u>not</u> available in the SOEP.

Variable Name	E11101\$\$
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	<p>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).</p> <p>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.</p> <p>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.</p> <p>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.</p> <p>There are still some cases where no annual work hours but individual labor income (I11101\$\$) can be observed. This finding can be explained by e.g. individuals who took maternity leave or did National Service.</p>
Format	<p>-1 = N/A - Child 0 = Not employed in the previous year The value of this variable ranges from 1 to 7457.</p> <p>The original survey variables provided below can be found in the file __P. This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions: D11101__ = Age of Individual</p>
Algorithm	<pre> if D11101\$\$ ge 16 then E11101\$\$=annual work hours imputation else E11101\$\$=0 </pre>

Variable Name	E11201\$\$
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 <u>not</u> available in the SOEP.

Variable Name	E11102\$\$
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	<p>-1 = N/A - Child 0 = Not Employed 1 = Employed</p> <p>This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions: D11101__ = Age of Individual E11101__ = Annual Work Hours of Individual I11110__ = Individual Labor Earnings</p>
Algorithm	<pre> if D11101\$\$ ge 16 then do if I11110\$\$ gt 0 and E11101\$\$ ge 52 then E11102\$\$=1 else E11102\$\$=0 end if D11101\$\$ lt 16 then E11102\$\$=-1 </pre>

Variable Name	E11103\$\$
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	<p>-1 = N/A - Child 1 = Full Time 2 = Part Time 3 = Not Working</p> <p>This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions: D11101__ = Age of Individual E11101__ = Annual Work Hours of Individual E11102__ = Employment Status of Individual</p>
Algorithm	<pre> if D11101\$\$ ge 16 then do 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 </pre>

Variable Name	E11104\$\$
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	<p>This variable is based on the individual's self-reported employment status at the time of the interview.</p> <p>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.</p> <p>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.</p>
Format	<p>-1 = N/A – Child -2 = Item-non response 1 = Working Now 2 = Not Working Now</p> <p>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</p>
Algorithm	
1984-1990:	<pre> 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) </pre>
1991-1995:	<pre> 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) </pre>
since 1996:	<pre> if D11101\$\$ ge 16 then do if VAR=1, 2, 3, 4, 8 then E11104\$\$=1 else if VAR=5, 6, 7, 9 then E11104\$\$=2 else E11104\$\$=-2 end else E11104\$\$=-1 (VAR=mp15, np11, op09, pp10, qp10, rp12, sp15, tp34, up09, vp10, wp07, xp13, yp19, zp09 \$\$=96-09) </pre>

Variable Name	E11105\$\$
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\$\$ = International standard classification of occupations). Occupation is coded as not applicable for individuals who were not working at the time of the interview.
Format	<p>-1 = N/A – Child -2 = Item Non-response</p> <p>A documentation for all other values of the ISCO-88 information (IS88\$\$ is a variable with four digits) can be found at: http://www.ilo.org</p> <p>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</p>
Algorithm	<pre> if X11103\$\$ = 1 then do if E11104\$\$ in (5,6,7,8) then E11105\$\$=0; else if E11104\$\$ in (1,2,3,4) and is88\$\$ le 0 then E11105\$\$=-1; else if E11104\$\$ in (1,2,3,4) and is88\$\$ gt 0 then E11105\$\$=IS88\$\$; else E11105\$\$=-2; end; </pre>

Variable Name	E11106\$\$
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	<p>-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 10 = Other</p> <p>The original survey variables provided below (NACE\$\$) can be found in the file _PGEN. This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions: E11104\$\$ = Primary Activity of Individual</p>
Algorithm	<pre> if X11103\$\$ = 1 then do 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; if nace\$\$ in (40,41) then E11106\$\$=2; if nace\$\$ in (10,11,12,13,14) then E11106\$\$=3; if nace\$\$ in (15,16,17,18,19,20,21,22,23,24,25,26,27,28,30,31,32,33, 37,96,97,100) then E11106\$\$=4; if nace\$\$ in (29,34,35,36,45) then E11106\$\$=5; if nace\$\$ in (50,51,52,55) then E11106\$\$=6; if nace\$\$ in (60,61,62,63,64) then E11106\$\$=7; if nace\$\$ in (65,66,67) then E11106\$\$=8; if nace\$\$ in (70,71,72,73,74,75,80,85,90,91,92,93,95,98,99) then E11106\$\$=9; end; else E11106\$\$=-2; end; </pre>

Variable Name	E11107\$\$
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 (NACE\$). 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
	-2 = Survey Non-response
	0 = Not Applicable
	1 = Agric.,Forestry
	2 = Fisheries
	3 = Energy/Water
	4 = Mining
	5 = Chemicals
	6 = Synthetics
	7 = Earth/Clay/Stone
	8 = Iron/Steel
	9 = Mechanical Eng
	10 = Electrical Eng
	11 = Wood/Paper/Print
	12 = Clothing/Text
	13 = Food Industry
	14 = Construction
	15 = Constr. Relate
	16 = Wholesale
	17 = Trading Agents
	18 = Retail
	19 = Train System
	20 = Postal System
	21 = Other Trans.
	22 = Financial Inst
	23 = Insurance
	24 = Restaurants
	25 = Service Indust
	26 = Trash Removal
	27 = Educ./Sport
	28 = Health Service
	29 = Legal Services
	30 = Other Services
	31 = Volunt./Church
	32 = Priv. Househld
	33 = Public Administration
	34 = Social Security
	99 = Not attributable

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	if X11103\$\$ = 1 then do
	if E11104\$\$ in (5,6,7,8) then E11107\$\$=0;
	else if E11104\$\$ in (1,2,3,4) and NACE\$\$ le 0 then E11107\$\$=-1;
	else if E11104\$\$ in (1,2,3,4) and NACE\$\$ gt 0 then do;
	if Nace\$\$ in (1,2) then E11107\$\$=1;
	if Nace\$\$ in (5) then E11107\$\$=2;
	if Nace\$\$ in (40,41) then E11107\$\$=3;
	if Nace\$\$ in (10,11,12,13,14) then E11107\$\$=4;
	if Nace\$\$ in (23,24) then E11107\$\$=5;
	if Nace\$\$ in (25) then E11107\$\$=6;
	if Nace\$\$ in (26) then E11107\$\$=7;
	if Nace\$\$ in (27,28) then E11107\$\$=8;
	if Nace\$\$ in (29,30,33) then E11107\$\$=9;
	if Nace\$\$ in (31,32) then E11107\$\$=10;
	if Nace\$\$ in (20,21,22) then E11107\$\$=11;
	if Nace\$\$ in (17,18,19) then E11107\$\$=12;
	if Nace\$\$ in (45) then E11107\$\$=14;
	if Nace\$\$ in (15,16) then E11107\$\$=13;
	if Nace\$\$ in (34,35,36) then E11107\$\$=15;
	if Nace\$\$ in (50,51,52) then E11107\$\$=16;

```

        if Nace$$ in (60,61,62,63,64)      then E11107$$=21;
        if Nace$$ in (65)                  then E11107$$=22;
        if Nace$$ in (66,67)               then E11107$$=23;
        if Nace$$ in (55)                  then E11107$$=24;
        if Nace$$ in (73,74)               then E11107$$=25;
        if Nace$$ in (37,95)               then E11107$$=26;
        if Nace$$ in (80,92)                then E11107$$=27;
        if Nace$$ in (85)                  then E11107$$=28;
        if NACE$$ in (70,71,72,93,98,99) then E11107$$=30;
        if NACE$$ in (91)                  then E11107$$=31;
        if NACE$$ in (90)                  then E11107$$=32;
        if NACE$$ in (75)                  then E11107$$=33;
        if NACE$$ in (96,97,100)           then E11107$$=99;
    end;
else E11107$$=-2;
end

```

Variable Name	H11101\$\$ Number of Household members age 0-14 H11102\$\$ Number of Household members age 15-18 H11103\$\$ Number of Household members age 0-1 H11104\$\$ Number of Household members age 2-4 H11105\$\$ Number of Household members age 5-7 H11106\$\$ Number of Household members age 8-10 H11107\$\$ Number of Household members age 11-12 H11108\$\$ Number of Household members age 13-15 H11109\$\$ Number of Household members age 16-18 H11110\$\$ Number of Household members age 19+ or 16-18 years old and independent
Unit of Observation	Household
Description	<p>These variables indicate the number of household members in the given age category living in the household at the time of the interview.</p> <p>H11109\$\$ 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\$\$ includes 16-18 year old youth who have completed Abitur or is in college, but exclude the head and the spouse. Only “residual” adults are counted in this variable.</p>
Method	These variables are the simple count of all individuals in the household whose age is in the listed category.
Format	<p>The value of this variable ranges from 0 to 20.</p> <p>The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<pre> *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 vp0902 wp0602 xp1202 yp1702 zp0802; array collg{*} 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{*} age\$\$; array ak102{*} marst\$\$; array ak103{*} insch\$\$; array ak104{*} collg\$\$; array ak105{*} hrel\$\$; array age14{*} age14\$\$; array age15_18{*} age18\$\$; array chld018{*} chld18\$\$; array age0_1{*} age1\$\$; array age2_4{*} age4\$\$; </pre>

```

array age5_7{*} age7$$;
array age8_10{*} age10$$;
array age11_12{*} age12$$;
array age13_15{*} age15$$;
array age16_18{*} age16$$;
array adults{*} adult$$;

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 scale ***;;
    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 <= age{i} < 19 then do;
      if (ak102{i} in (2,3,4,5,6,M) & ak103{i}= 1 &
        ak104{i}= 0) or (ak102{i} = 5 & ak103{i}= 1 &
        ak104{i}= 0 & ak105{i}= 3) or
        (ad100{i} = 20) then age16_18{i}=1;
      if ak105{i} in (1,2) then age16_18{i}=0;
    end;

    if age16_18{i}=0 & (16<=age{i} < 19) then adults{i}=1;
    if age{i} lt 0 then adults{i}=1;
    if ak105{i} in (1,2) then adults{i}=0;
    if ak102{i}=1 & (16 <= age{i} < 19) then adults{i}=0;
    if age16_18{i}=1 then adults{i}=0;
  end;
end;

*** All variables are then summed by household id number (X11102$$)***

```

Variable Names	H11112\$\$
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	<p>0 = Not present 1 = Present</p> <p>The variables provided below can be found in the \$PEQUIV files. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<pre> if X11103\$\$=1 then do; H11112\$\$=0; if D11105\$\$=2 then H11112\$\$=1; end; </pre>

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 (referred to below as W11114\$\$)
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 (referred to below as W11115\$\$)
Unit of Observation	Household
Description	Scale used by Organization for Economic Cooperation and Development (1982), see also Hagenaars 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 (e.g. Square root of the Household size referred to below as W11116\$\$)
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: $EI = D/S^e$ Where equivalent income (EI) 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	L11101\$\$
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	<p>-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-Wuerttemberg 9 = Bavaria 11 = Berlin (East) 12 = Mecklenburg-Vorpommern 13 = Brandenburg 14 = Saxony-Anhalt 15 = Thuringen 16 = Saxony</p> <p>These states can be collapsed into regions. From 1984 through 1989 three regions can be defined to include the following states:</p> <p>North: Berlin (0), Schleswig-Holstein (1), Hamburg (2), Lower-Saxony (3), Bremen (4) South: Hessen (6), Baden-Wuerttemberg (8), Bavaria (9) West North-Rhine-Westfalia (5), Rheinland-Pfalz, Saarland (7)</p> <p>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-Wuerttemberg (8), Bavaria (9) West: North-Rhine-Westfalia (5), Rheinland-Pfalz, Saarland (7) East: Berlin (0), Berlin East (11), Mecklenburg-Vorpommern (12), Brandenburg (13) Saxony-Anhalt (14), Thuringen (15), Saxony (16)</p> <p>This algorithm omits individuals with survey non-responses. Original variables below can be found in _HBRUTTO files</p>
Algorithm	L11101\$\$=Ybula

Variable Name	L11102\$\$
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 1 = West Germany 2 = East Germany
Algorithm	This algorithm omits individuals with survey non-responses.
	Original variables below can be found in _HBRUTTO files
	L11102\$\$=1 (\$\$=84-90) L11102\$\$=Ysampreg (\$\$=90, ...)

Variable Name	Y11101\$\$
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 2005 (survey year 2006).
Method	<p>This value of this variable is derived from the “DESTATIS 2009, Preise. Verbraucherpreisindex für Deutschland. Lange Reihen ab 1948- Juni 2009.”</p> <p>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.</p> <p>Example: $I1110385 * (Y1110192/Y1110185)$</p>
Format	N/A

Price Index for West German States

1984:	Y1110184 = 66,8
1985:	Y1110185 = 68,5
1986:	Y1110186 = 69,9
1987:	Y1110187 = 69,8
1988:	Y1110188 = 69,9
1989:	Y1110189 = 70,8
1990:	Y1110190 = 72,8
1991:	Y1110191 = 74,7
1992:	Y1110192 = 77,5
1993:	Y1110193 = 80,6
1994:	Y1110194 = 83,4
1995:	Y1110195 = 85,7
1996:	Y1110196 = 87,1
1997:	Y1110197 = 88,2
1998:	Y1110198 = 89,9
1999:	Y1110199 = 90,7
2000:	Y1110100 = 91,3

Price Index for East German States

N/A
N/A
N/A
N/A
N/A
N/A
N/A
N/A
Y1110192 = 65,8
Y1110193 = 74,6
Y1110194 = 82,5
Y1110195 = 85,4
Y1110196 = 87,1
Y1110197 = 88,8
Y1110198 = 90,8
Y1110199 = 91,7
Y1110100 = 92,1

Price Index for German States

2001:	Y1110101 = 92,7
2002:	Y1110102 = 94,5
2003:	Y1110103 = 95,9
2004:	Y1110104 = 96,9
2005:	Y1110105 = 98,5
2006:	Y1110106 = 100,0
2007:	Y1110107 = 101,6
2008:	Y1110108 = 103,9
2009:	Y1110109 = 106,6

Variable Name	I11101\$\$
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-employment 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	<p>The value of this variable ranges from 0 to 5.000.000. This variable is in current year EURO.</p> <p>This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions:</p> <p style="padding-left: 40px;">I11103__ = Household Labor Earnings I11104__ = Household Asset Income I11106__ = Household Private Transfers I11117__ = Household Private Retirement Income</p>
Algorithm	$I11101\$ = I11103\$ + I11104\$ + I11106\$ + I11117\$$

Variable Name	I11102\$\$
Variable Label	Household Post-Government Income
Unit of Observation	Household
Period	Annual
Description	This variable represents the combined income after taxes and government transfers in the previous year of all individuals in the household.
Method	<p>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 allowances, 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).</p> <p>Household post-government income has no negative values.</p>
Format	<p>The value of this variable ranges from 0 to 5.000.000. This variable is in current year EURO.</p> <p>This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions:</p> <ul style="list-style-type: none"> 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	$I11102\$ = I11103\$ + I11104\$ + I11106\$ + I11107\$ + I11108\$ + I11117\$ - I11109\$$

Variable Name	I11103\$\$
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	<p>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.</p> <p>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. Since 1991 indemnity payments and commuting expenses or travel grants are also considered.</p>
Format	<p>The value of this variable ranges from 0 to 5.000.000.</p> <p>This variable is in current year EURO.</p> <p>This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions: I11110__ = Individual Labor Earnings</p>
Algorithm	$I11103\$ = \text{sum of } I11110\$ \text{ over all individuals in the household}$

Variable Name	I11104\$\$
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	<p>Asset flows include income from interest, dividends, and rent.</p> <p>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:</p> <ul style="list-style-type: none"> 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 <p>Starting in year 2001 (wave R) an additional item was offered:</p> <ul style="list-style-type: none"> 10,000 to 20,000 DM 20,000 DM and over <p>Since year 2002 (wave S) all items are asked for Euro:</p> <ul style="list-style-type: none"> 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 <p>These respondents are assigned an interest and dividend amount based on uniformly distributed random numbers within their income range.</p> <p>Rental income is the amount of income from rent minus any operation and maintenance costs.</p> <p>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.</p>
Format	<p>The value of this variable ranges from 0 to 5.000.000.</p> <p>This variable is in current year EURO.</p> <p>The survey variables provided below are part of the \$PEQUIV-file.</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	$I11104\$ = (RENTY\$ - OPERY\$) + DIVDY\$$

Variable Name	I11105\$\$
Variable Label	Household Imputed Rental Value
Unit of Observation	Household
Period	Annual
Description	This variable represents the imputed rental value of owner occupied housing and for renters with below markets rent.
Method	<p>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).</p> <p>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 <u>and</u> individuals living in households with reduced rent such as employer provided flats, social housing or rent-free households. The resulting estimate represents a gross value at market prices (without costs for heating and warm water). For owner-occupiers owner-specific costs for taxation, maintenance and operating costs as well as interest on mortgages were deducted yielding a net value which can be interpreted as the appropriate income advantage of owner-occupied housing. For rent-free households and persons living in households with below market rents no further deductions have to be made.</p> <p>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/m² per month is used for West Germany in 1988 and rises to 5.52 DM/m² per month in 1998. The average interest burden in East Germany was slightly lower, at 4.14 DM/m² per month in 1998.</p> <p>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.</p> <p>For further details see: Frick and Grabka (2001) and Frick and Grabka (2003):</p>
Format	The value of this variable ranges from 0 to 5.000.000. This variable is in current year EURO.
Algorithm	N/a

Variable Name	I11106\$\$
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 this income was not specifically identified. Starting in wave R an additional question identifies alimony separately (variable \$p2o03 in SOEP file \$PKAL: \$ = R, S, ...). 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. 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	$I11106\$ = \text{sum of } (I11106\$ + I11106\$) \text{ over all individuals in the household}$

Variable Name	I11107\$\$
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	<p>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, support for special circumstances, social assistance for elderly and unemployment benefit II.</p> <p>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.</p> <p>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.</p> <p>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. In 2005 social assistance for elderly was asked the first time. In 2006 unemployment benefit II was asked the first time and replaced unemployment assistance. Since 2009 additional child benefit was asked the first time.</p>
Format	<p>The value of this variable ranges from 0 to 5.000.000.</p> <p>This variable is in current year EURO.</p> <p>The survey variables provided below are part of the \$PEQUIV-file.</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	$I1110704 = [\text{sum of } (IUNBY\$ + IUNAY\$ + ISUBY\$ + IERET\$ + IMATY\$ + ISTUY\$) \text{ over all individuals in the household}] + HOUSE\$ + CHSPT\$ + NURSH\$ + SUBST\$ + SPHLP\$ + HSUP\$ + SSOLD\$ + ALG2\$ + ADCHB\$$

Variable Name	I11108\$\$
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	<p>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).</p> <p>In 1993 through 1994 pension income from East German pensions (\$p7902o and \$p7912o) is assigned to other pension income.</p> <p>In case of partial unit-non responding households this information has been imputed. For details see: Frick, Grabka & Groh-Samberg (2010).</p>
Format	<p>The value of this variable ranges from 0 to 5.000.000.</p> <p>This variable is in current year EURO.</p> <p>The survey variables provided below are part of the \$PEQUIV-file.</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	
1984-1985, 2002:	$I11108\$ = \text{sum of } (IOLDY\$ + IWIDY\$ + ICOMP\$ + IPRVP\$) \text{ over all individuals in the household}$ <p style="text-align: right;">(\$\$=84-85, 02)</p>
1986-2001, since 2003:	$I11108\$ = \text{sum of } (igrv1\$ + igrv2\$ + ismp1\$ + ismp2\$ + iciv1\$ + iciv2\$ + iwar1\$ + iwar2\$ + iagr1\$ + iagr2\$ + iguv1\$ + iguv2\$) \text{ over all individuals in the household}$ <p style="text-align: right;">(\$\$=86-01,03,...)</p>

Variable Name	I11109\$\$
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	<p>The value of this variable ranges from 0 to 5.000.000.</p> <p>This variable is in current year EURO.</p> <p>The survey variables provided below are part of the \$PEQUIV-file.</p>
Algorithm	$I11109\$ = I11111\$ + I11112\$$

Variable Name	I11110\$\$
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	<p>Labor earnings include wages and salary from all employment including training, primary and secondary jobs, and self-employment, plus income from bonuses, over-time, and profit-sharing.</p> <p>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.</p> <p>Since 1991 indemnity payments, since 1996 military service payments and since 2006 commuting expenses or travel grants are also considered.</p> <p>In case of partial unit-non responding households this information has been imputed. For details see: Frick, Grabka & Groh-Samberg (2010).</p>
Format	<p>The value of this variable ranges from 0 to 5.000.000.</p> <p>This variable is in current year EURO.</p> <p>The survey variables provided below are part of the \$PEQUIV-file.</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	$I11110\$ = IJOB1\$ + IJOB2\$ + ISELF\$ + IMILT\$ + I13LY\$ + I14LY\$ +$ $IXMAS\$ + IHOLY\$ + IGRAY\$ + IOTHY\$ + IDEMY\$ +$ $ITRAY\$$

Variable Name	I11111\$\$
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. This variable is in current year EURO.
Algorithm	N/a

Variable Name	I11112\$\$
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. This variable is in current year EURO.
Algorithm	N/a

Variable Name	I11113\$\$	Household Post-Government Income (TAXSIM)
	I11114\$\$	Total Household Taxes (TAXSIM)
	I11115\$\$	Household State Taxes (TAXSIM)
	I11116\$\$	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	I11117\$\$
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	<p>The value of this variable ranges from 0 to 5.000.000. This variable is in current year EURO.</p> <p>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.</p>
Algorithm	
1984-1985:	N/a
2002-2003:	$I11117\$ = \text{sum of (ICOMP\$ + IPRVP\$) over all individuals in the household}$ <p style="text-align: right;">(\$\$= 02-03)</p>
1986-2001, since 2004:	$I11117\$ = \text{sum of (ivbl1\$ + ivbl2\$ + icom1\$ + icom2\$ + iprv1\$ + iprv2\$ + ison1\$ + ison2\$) over all individuals in the household}$ <p style="text-align: right;">(\$\$= 86-01, 04, ...)</p>

Variable Name	I11118\$\$
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), 2500 Euro (wave S-U) or 500 Euro (since wave V). 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.
Algorithm	<p>This algorithm omits individuals with survey non-responses. The original survey variables provided below can be found in the file _H.</p> <p>I11118\$\$ = Ywind</p> <p>Windfall income variable list by survey year - each entry denoted in algorithm as Ywind: QH4505 RH4505 SH4505 TH44 UH44 VH47 WH47 XH47 YH48 ZH48</p>

Variable Name	I11201\$\$
Variable Label	Impute Household Pre-Government Income
Unit of Observation	Household
Description	<p>This variable indicates the percentage share of income that has been imputed for Household Pre-Government Income (I11101\$).</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 and -3 only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	I11201\$\$= percentage share of income that has been imputed for I11101\$

Variable Name	I11202\$\$
Variable Label	Impute Household Post-Government Income
Unit of Observation	Household
Description	<p>This variable indicates the percentage share of income that has been imputed for household post-government income (I11102\$\$).</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 and -3 only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	I11202\$\$= percentage share of income that has been imputed for I11102\$\$

Variable Name	I11203\$\$
Variable Label	Impute Household Labor Income
Unit of Observation	Household
Description	<p>This variable indicates the percentage share of income that has been imputed for household labor income (I11103\$\$).</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	I11203\$\$= percentage share of income that has been imputed for I11103\$\$

Variable Name	I11204\$\$
Variable Label	Impute Household Asset Income
Unit of Observation	Household
Description	<p>This variable indicates the percentage share of income that has been imputed for household asset income (I11104\$\$).</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions: I11104__ = Household Asset Income</p>
Algorithm	I11204\$\$= percentage share of income that has been imputed for I11104\$\$

Variable Name	I11205\$\$
Variable Label	Impute Household Imputed Rental Value
Unit of Observation	Household
Description	This variable indicates if the imputed rental value (I11105\$\$) 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 1 = Fully Imputed This algorithm omits individuals with survey non-responses.
Algorithm	$I11205\$ = 1 \text{ if } I11105\$ > 0, \text{ else } I11205\$ = 0.$

Variable Name	I11206\$\$
Variable Label	Impute Household Private Transfers
Unit of Observation	Household
Description	<p>This variable indicates the percentage share of income that has been imputed for household private transfers (I11106\$\$).</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	I11206\$\$= percentage share of income that has been imputed for I11106\$\$

Variable Name	I11207\$\$
Variable Label	Impute Household Public Transfers
Unit of Observation	Household
Description	<p>This variable indicates the percentage share of income that has been imputed for household public transfers (I11107\$\$).</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	I11207\$\$= percentage share of income that has been imputed for I11107\$\$

Variable Name	I11208\$\$
Variable Label	Impute Household Social Security Pensions
Unit of Observation	Household
Description	<p>This variable indicates the percentage share of income that has been imputed for household social security pensions (I11108\$).</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	I11208\$\$= percentage share of income that has been imputed for I11108\$\$

Variable Name	I11209\$\$
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 1 = Fully Imputed This algorithm omits individuals with survey non-responses. Equivalent Data File Variable Definitions: I11109__ = Total Household Taxes
Algorithm	$I11209\$ = 1 \text{ if } I11109\$ > 0, \text{ else } I11209\$ = 0.$

Variable Name	I11210\$\$
Variable Label	Impute Individual Labor Earnings
Unit of Observation	Individual
Description	<p>This variable indicates the percentage share of income that has been imputed for individual labor earnings (I11110\$\$).</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	I11210\$\$= percentage share of income that has been imputed for I11110\$\$

Variable Name	I11217\$\$
Variable Label	Impute Household Private Retirement Income
Unit of Observation	Household
Description	<p>This variable indicates the percentage share of income that has been imputed for household private retirement income (I11117\$).</p> <p>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)</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <p>-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</p> <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	I11217\$\$= percentage share of income that has been imputed for I11117\$\$

Variable Name	I11218\$\$
Variable Label	Impute Household Windfall Income
Unit of Observation	Household
Description	<p>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)</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <p>-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</p> <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	N/a

Variable Name	RENTY\$\$
Variable Label	Income from rental and leasing
Unit of Observation	Household
Period	Annual
Description	<p>This variable represents the household income from rental and leasing in the last year.</p> <p>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.</p>
Method	Transcribed variable.
Format	<p>The value of this variable ranges from 0 to 999.999.</p> <p>This variable is in current year EURO.</p> <p>The original survey variables provided below can be found in the file _H. This algorithm omits individuals with survey non-responses.</p>

Algorithm

```

array cm101{*} ah41 bh35 ch47 dh47 eh38 fh38 gh38 hhrenty ih41 jh41 kh41 lh41 mh41 nh41
oh41 ph41 qh41 rh41 sh41 th39 uh39 vh38 wh38 xh38 yh39 zh39;
*** imputed values due to item-non response ***
array cm102{*} xah41 xbh35 xch47 xdh47 xeh38 xfh38 xgh38 out xih41 xjh41 xkh41 xlh41
mh41 xnh41 xoh41 xph41 xqh41 xrh41 xsh41 xth39 xuh39 xv38 xwh38
xxh38 xyh39 xzh39;
array cm103{*} temp1$$;
array cm104{*} renty$$;

do i = 1 to dim(netto);
  cm103{i}=.;
  if netto{i} >= 10 & < 20 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	OPERY\$\$
Variable Label	Operation, maintenance costs
Unit of Observation	Household
Period	Annual
Description	<p>This variable represents the household operation and maintenance costs in the last year.</p> <p>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.</p>
Method	Transcribed variable.
Format	<p>The value of this variable ranges from 0 to 999.999.</p> <p>This variable is in current year EURO.</p> <p>The original survey variables provided below can be found in the file _H. This algorithm omits individuals with survey non-responses.</p>

Algorithm

```

array ct101{*} ah4201 bh3601 ch4801 dh4801 eh3901 fh3901 gh3901 hhopery ih4201 jh4201
kh4201 lh4201 mh4201 nh4201 oh4201 ph4201 qh4201 rh4201 sh4201 th4001
uh4001 vh3901 wh3901 xh3901 yh4001 zh4001;
*** 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 xvh3901 xwh3901 xxh3901 xyh4001
xzh4001;
array ct103{*} temp$$
array ct104{*} opery$$;

do i = 1 to dim(netto);
  ct103{i}=.;
  if netto{i} >= 10 & < 20) 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	DIVDY\$\$
Variable Label	Interest, dividend income	
Unit of Observation	Household	
Period	Annual	
Description	This variable represents the household income from interest and dividends in the last year.	
Method	<p>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:</p> <p>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</p> <p>Starting in year 2001 (wave R) an additional item was offered: 10,000 to 20,000 DM 20,000 DM and over</p> <p>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</p> <p>These respondents are assigned an interest and dividend amount based on uniformly distributed random numbers within their income range.</p>	
Format	<p>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.</p>	

Algorithm

```

array cx101{*} ah45 bh3801 ch5001 dh5001 eh4101 fh4101 gh4101 hh4701 ih4401 jh4401
kh4401 lh4401 mh4401 nh4401 oh4401 ph4401 qh4401 rh4401 sh4401 th4201
uh4201 vh4501 wh4501 xh4501 yh4601 zh4601;
*** 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 xvh4501 xwh4501 xxh4501 xyh4601;
xzh4601
array cx103{*} temp$$;
array cx104{*} divdy$$;

if ah45=.B then ah45=0;
do i = 1 to dim(netto);
cx103{i}=.;
if netto{i} >= 10 & < 20 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	CHSPT\$\$
Variable Label	Child allowance
Unit of Observation	Household
Period	Annual
Description	This variable represents the household income from child allowances in the last year.
Method	<p>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.</p> <p>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.</p> <p>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.</p>
Format	<p>The value of this variable ranges from 0 to 999.999.</p> <p>This variable is in current year EURO.</p> <p>The original survey variables provided below can be found in the file _H.</p> <p>This algorithm omits individuals with survey non-responses.</p>

Algorithm

```

array dp101{*} out bh3303 ch4503 dh4503 eh3603 fh3603 gh3603 hh4503 ih4603 jh4603
kh4603 lh4603 mh4603 nh4603 oh4603 ph4603 qh50 rh4603 sh4603 th4503
uh4503 vh4803 wh4803 xh4803 yh4903 zh4903;
*** 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 xvh4803 xwh4803 xxh4803 xyh4903
xzh4903;
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{*} temp$$;
array dp106{*} out out out out out out out out out out out out out out out out
sh4602 th4502 uh4502 vh4802 wh4802 xh4802 yh4902 zh4902;
array dp105{*} chspt$$;

do i = 1 to dim(netto);
dp104{i}=.;
if netto{i} >= 10 & < 20 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	HOUSE\$\$
Variable Label	Housing allowance
Unit of Observation	Household
Period	Annual
Description	This variable represents the household income from housing allowance in the last year.
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

```

array dl101{*} ah29 bh2802 ch4002 dh4002 eh3102 fh3102 gh3102 hh4002 ih4502 jh4502
               kh4502 lh4502 mh4502 nh4502 oh4502 ph4502 qh47 rh4605 sh4605 th4505
               uh4505 vh4805 wh4808 xh4808 yh4908 zh4911;
array dl102{*} ah30 bh2803 ch4003 dh4003 eh3103 fh3103 gh3103 hh4003 ih4503 jh4503
               kh4503 lh4503 mh4503 nh4503 oh4503 ph4503 qh48 rh4606 sh4606 th4506
               uh4506 vh4806 wh4809 xh4809 yh4909 zh4912;
*** 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 rh4606
               xsh4606 xth4506 xuh4506 xv4806 xwh4809 xxh4809 xyh4909 xzh4912;
array dl104{*} temp$$4
array dl105{*} temp$$;
array dl106{*} house$$;

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} >= 10 & < 20 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	NURSH\$\$
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 : N/a since 2001: array dzc101{ *} rh4609 sh4609 th4509 uh4509 vh4809 wh4812 xh4812 yh4912 zh4915; *** imputed values due to item-non response *** array dzc102{ *} xrh4609 xsh4609 xth4509 xuh4509 xv4809 xwh4812 xxh4812 xyh4912 xzh4915; array dzc103{ *} rh4608 sh4608 th4508 uh4508 vh4808 wh4811 xh4811 yh4911 zh4914; array dzc104{ *} nursh\$\$; 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	SUBST\$\$
Variable Label	Social assistance
Unit of Observation	Household
Period	Annual
Description	This variable represents the household income from social assistance in the last year.
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 <u>and</u> 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	<pre> array dt101{*} ah34 bh3002 ch4202 dh4202 eh3302 fh3302 gh3302 hh4202 out out out out out out out out out rh4702 sh4702 th4602 uh4602 vh4902 wh4902 xh4902 yh5002 zh5002; array dt102{*} ah35 bh3003 ch4203 dh4203 eh3303 fh3303 gh3303 hh4203 out out out out out out out out out rh4703 sh4703 th4603 uh4603 vh4903 wh4903 xh4903 yh5003 zh5003; *** 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 xrh4703 xsh4703 xth4603 xuh4603 xvh4903 xwh4903 xxh4903 xyh5003 xzh5003; *** 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 out out out out; array dt105{*} temp1\$\$ array dt106{*} temp2\$\$; array dt107{*} subst\$\$; do i = 1 to dim(netto); dt105{i}=.; dt106{i}=.; if netto{i} >= 10 & < 20 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; </pre>

Variable Name	SPHLP\$\$
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 in the last year.
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 _H. This algorithm omits individuals with survey non-responses.
Algorithm	<pre> array dx101{*} ah3601 bh3102 ch4302 dh4302 eh3402 fh3402 gh3402 hh4302 out out out out out out out out out rh4705 sh4705 th4605 uh4605 vh4908 wh4908 xh4908 yh5008 zh5008; array dx102{*} ah37 bh3103 ch4303 dh4303 eh3403 fh3403 gh3403 hh4303 out out out out out out out out out rh4706 sh4706 th4606 uh4606 vh4909 wh4909 xh4909 yh5009 zh5009; *** 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 xrh4706 xsh4706 xth4606 xuh4606 xvh4909 xwh4909 xxh4909 xyh5009 xzh5009; array dx104{*} temp1\$\$; array dx105{*} temp2\$\$; array dx106{*} sphlp\$\$; do i = 1 to dim(netto); dx104{i}=.; dx105{i}=.; if netto{i} >= 10 & < 20 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; </pre>

Variable Name	SSOLD\$\$
Variable Label	Social assistance for elderly (Grundsicherung im Alter)
Unit of Observation	Household
Period	Annual
Description	This variable represents the household income from Social assistance for elderly in the last year.
Method	Social assistance for elderly is the product of the number of months that benefit was claimed in the previous year and the average amount per month. It was asked the first time in wave V (variable vh4906 in SOEP file VH).
Format	<p>The value of this variable ranges from 0 to 99.999. This variable is in current year EURO.</p> <p>The original survey variables provided below can be found in the file _H. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<pre> array dy101{*} vh4905 wh4905 xh4905 yh5005 zh5005; array dy102{*} vh4906 wh4906 xh4906 yh5006 zh5006; array dy103{*} xvh4906 xwh4906 xxh4906 xyh5006 xzh5006; array dy104{*} temp1\$\$; array dy105{*} temp2\$\$; array dy106{*} ssold\$\$; do i = 1 to dim(dy101); dy104{i}=.; dy105{i}=.; if aa100{i}=1 then do; if dy103{i} lt 0 then dy103{i}=0; if dy101{i} eq .B then dy104{i}=0; if dy101{i} in (.A,.C) then dy104{i}=11; if dy101{i} ge 0 then dy104{i}=dy101{i}; if dy102{i} in (.A,.C) then do; if dy101{i} le 0 and dy103{i} gt 0 then dy104{i}=11; dy105{i}=dy103{i}; end; else if dy102{i} eq .B then dy105{i}=0; else dy105{i}=dy102{i}; dy106{i}=dy104{i}*dy105{i}; end; else dy106{i}=.; end; </pre>

Variable Name	ALG2\$\$
Variable Label	Unemployment benefit II
Unit of Observation	Household
Period	Annual
Description	This variable represents the household income from unemployment benefit II including social benefit and child allowance in the last year.
Method	Unemployment benefit II is the product of the number of months that benefit was claimed in the previous year and the average amount per month. It was asked the first time in wave W (variable wh4806 in SOEP file WH).
Format	<p>The value of this variable ranges from 0 to 99.999. This variable is in current year EURO.</p> <p>The original survey variables provided below can be found in the file _H. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<pre> array ey101{*} wh4805 xh4805 yh4905 zh4908; array ey102{*} wh4806 xh4806 yh4906 zh4909; array ey103{*} xwh4806 xxh4806 xyh4906 xzh4909; array ey104{*} temp1\$\$; array ey105{*} temp2\$\$; array ey106{*} alg2\$\$; do i = 1 to dim(ey101); ey104{i}=.; ey105{i}=.; if \$netto >= 10 & < 20 then do; if ey103{i} lt 0 then ey103{i}=0; if ey101{i} eq .B then ey104{i}=0; if ey101{i} in (.A,.C) then ey104{i}=11; if ey101{i} ge 0 then ey104{i}=ey101{i}; if ey102{i} in (.A,.C) then do; if ey101{i} le 0 and ey103{i} gt 0 then ey104{i}=11; ey105{i}=ey103{i}; end; else if ey102{i} eq .B then ey105{i}=0; else ey105{i}=ey102{i}; ey106{i}=ey104{i}*ey105{i}; end; else ey106{i}=.S; end; </pre>

Variable Name	HSUP\$\$
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 in the last year.
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

```

array bzc101{*} misses misses misses misses misses misses misses misses misses misses
               misses misses ms3904 ns3904 os3904 ps3904 qh3904 rh3904 sh3904 th3504
               uh3504 vh3602 wh3602 xh3602 yh3702 zh3702;
*** 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 xvh3602 xwh3602 xxh3602 xyh3702 xzh3702;
array bzc103{*} temp1$$;
array bzc106{*} hsup$$;

do i = 1 to dim(netto);
  bzc103{i}=.;
  if netto{i} >= 10 & < 20 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	LOSSR\$\$
Variable Label	Losses from renting and leasing
Unit of Observation	Household
Period	Annual
Description	This variable represents the household tax deductible costs or losses from renting and leasing incurred in the last.
Method	Transcribed variable. It was asked the first time in wave V (variable vh4002 in SOEP file VH).
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

```

array cw101{*} vh4002 wh4002 xh4002 yh4102 zh4102;
array cw102{*} xv4002 xwh4002 xxh4002 xyh4102 xzh4102;
array cw103{*} temp1$$;
array cw104{*} lossr$$;

do i = 1 to dim(netto);
  cw103{i}=.;
  if netto{i} >= 10 & < 20 then do;
    if cw102{i} lt 0 then cw102{i} = 0;
    if cw101{i}=.A or cw101{i}=.C then cw103{i}=cw102{i};
    else if cw101{i}=.B then cw103{i}=0;
    else cw103{i}=cw101{i};
    cw104{i}=cw103{i};
  end;
  else cw104{i}=.S;
end;

```

Variable Name	LOSSC\$\$
Variable Label	Losses from capital investment
Unit of Observation	Household
Period	Annual
Description	This variable represents the household losses from capital investment in the last year.
Method	Transcribed variable. It was asked the first time in wave V (variable vh4409 in SOEP file VH).
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

```

array cw201{*} vh4409 wh4409 xh4409 yh4509 zh4509;
array cw202{*} xvh4409 xwh4409 xxh4409 xyh4509 xzh4509;
array cw203{*} temp1$$;
array cw204{*} lossc$$;

do i = 1 to dim(netto);
  cw203{i}=.;
  if netto{i} >= 10 & < 20 then do;
    if cw202{i} lt 0 then cw202{i} = 0;
    if cw201{i} in (.A.,C) then cw203{i}=cw202{i};
    else if cw201{i}=B then cw203{i}=0;
    else cw203{i}=cw201{i};
    cw204{i}=cw203{i};
  end;
  else cw204{i}=S;
end;

```

Variable Name	ADCHB\$\$
Variable Label	Additional Child Benefit
Unit of Observation	Household
Period	Annual
Description	This variable represents the household income from additional child benefit in the last year.
Method	Transcribed variable. It was asked the first time in wave Z (variable ZH4906 in SOEP file ZH).
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

```

array kp101{*} zh4906;
array kp102{*} zh5103;
array kp103{*} zh4905;
array kp104{*} xzh4906;
array kp105{*} adchb09;

do i = 1 to dim(netto);
  if netto=1 then do;
    kp104{i}=0;
    if kp102{i} in (.A.,C) then kp102{i}=1;      *** min. 1 Kind ***;
    if kp102{i} in (.B.,) then kp102{i}=0;
    if kp101{i} in (.A.,C) then kp101{i}=140*kp102{i};  *** INR --> Maximum ***;
    if kp103{i} in (.B.,) then kp103{i}=0;
    if kp103{i} in (.A.,C) then kp103{i}=6;          *** INR --> 6 Monate ***;

    if (i ge 1 and i le 25) then kp105{i}=0;      *** wave A-Y *;
    else if (i ge 26)      then kp105{i}=kp101{i}*kp103{i};
  end;
  else kp105{i}=.S;
end;

```

Variable Name	FRENTY\$\$
Variable Label	Impute Income from rental and leasing
Unit of Observation	Household
Description	<p>This variable indicates if income from rental and leasing (RENTY\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FOPERY\$\$
Variable Label	Impute Operation, maintenance costs
Unit of Observation	Household
Description	<p>This variable indicates if operation or maintenance costs (OPERY\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed</p> <p>1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FDIVDY\$\$
Variable Label	Impute Interest, dividend income
Unit of Observation	Household
Description	<p>This variable indicates if income from interest or dividends (DIVDY\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FCHSPT\$\$
Variable Label	Impute Child allowance
Unit of Observation	Household
Description	<p>This variable indicates if income from child allowances (CHSPT\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FHOUSE\$\$
Variable Label	Impute Housing benefit
Unit of Observation	Household
Description	<p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FNURSH\$\$
Variable Label	Impute Nursing allowances
Unit of Observation	Household
Description	<p>This variable indicates if income from nursing allowances (NURSH\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed</p> <p>1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FSUBST\$\$
Variable Label	Impute Social assistance
Unit of Observation	Household
Description	<p>This variable indicates if income from social assistance (SUBST\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FSPHLP\$\$
Variable Label	Impute Social assistance for special circumstances
Unit of Observation	Household
Description	<p>This variable indicates if income from social assistance for special circumstances (SPHLP\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FSSOLD\$
Variable Label	Impute Social assistance for elderly
Unit of Observation	Household
Description	<p>This variable indicates if income from social assistance for elderly (SSOLD\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FALG2\$\$
Variable Label	Imputation flag: Unemployment benefit II
Unit of Observation	Household
Description	<p>This variable indicates if income from unemployment benefit II (ALG2\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FHSUP\$\$
Variable Label	Impute Housing support for owner-occupiers
Unit of Observation	Household
Description	<p>This variable indicates if income from housing support for owner-occupiers (HSUP\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FLOSSR\$
Variable Label	Impute Losses from renting and leasing
Unit of Observation	Household
Description	<p>This variable indicates if losses from renting and leasing (LOSSR\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FLOSSC\$
Variable Label	Impute Losses from capital investment
Unit of Observation	Household
Description	<p>This variable indicates if losses from capital investment (LOSSC\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FADCHB\$
Variable Label	Impute additional child benefit
Unit of Observation	Household
Description	<p>This variable indicates if income from additional child benefit (ADCHB\$\$) has been imputed. In case of item-non-response this information is imputed by applying the maximum available amount (140 €per month per child).</p> <p>In 2009 about 188 household stated that they received additional child benefit in the previous year. However, after various sensitivity checks 159 households were set to “-2”. These households either did not receive any child benefit in the previous year – which is a prerequisite for eligibility – nor had incomes which made the household eligible for that public transfer.</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	IJOB1\$\$
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. 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	$IJOB1\$ = (\$P2A02 * \$P2A03)$

Variable Name	IJOB2\$\$
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	$IJOB2\$ = (\$P2C02 * \$P2C03)$

Variable Name	ISELF\$\$
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. 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	$\text{ISELF}\$ = (\$P2B02 * \$P2B03)$

Variable Name	IOLDY\$\$
Variable Label	Combined old-age, disability and civil servants pensions
Unit of Observation	Individual
Period	Annual
Description	<p>This variable represents income from combined old-age, disability and civil servants pensions of individuals in the household 16 years of age and older.</p> <p>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.</p>
Method	<p>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.</p> <p>If the information about the number of months is missing, the sample mean of that variable has been assigned.</p>
Format	<p>The value of this variable ranges from 0 to 999.999.</p> <p>This variable is in current year EURO.</p> <p>The original survey variables provided below can be found in the file _PKAL.</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	$\text{IOLDY}\$ = (\text{\$P2D02} * \text{\$P2D03})$

Variable Name	IWIDY\$\$
Variable Label	Combined widows and orphans pension
Unit of Observation	Individual
Period	Annual
Description	<p>This variable represents income from combined widows and orphans pension of individuals in the household 16 years of age and older.</p> <p>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.</p>
Method	<p>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.</p> <p>If the information about the number of months is missing, the sample mean of that variable has been assigned.</p>
Format	<p>The value of this variable ranges from 0 to 999.999.</p> <p>This variable is in current year EURO.</p> <p>The original survey variables provided below can be found in the file _PKAL.</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	$IWIDY\$ = (\$P2E02 * \$P2E03)$

Variable Name	ICOMP\$\$
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. 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	1984-2001: N/a 2002-2003: $ICOMP\$ = (\$P2P02 * \$P2P03)$ (\$\$=02 – 03, \$ = S-T) since 2004: N/a

Variable Name	IPRVP\$\$
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. 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	1984-2001: N/a 2002-2003: $IPRVP\$ = (\$P2Q02 * \$P2Q03)$ ($\$ = 02 - 03$, $\$ = S-T$) since 2004: N/a

Variable Name	IUNBY\$\$
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	<p>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.</p> <p>If the information about the number of months is missing, the sample mean of that variable has been assigned.</p> <p>In case of partial unit-non responding households this information has been imputed.</p> <p>For details see: Frick, Grabka & Groh-Samberg (2010).</p>
Format	<p>The value of this variable ranges from 0 to 999.999.</p> <p>This variable is in current year EURO.</p> <p>The original survey variables provided below can be found in the file _PKAL.</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	$IUNBY\$ = (\$P2F02 * \$P2F03)$

Variable Name	IUNAY\$\$
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. In 2006 this information is no longer relevant, given that unemployment assistance has been replaced by unemployment benefit II (ALG2\$\$)
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. 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	1984-2005: IUNAY\$\$ = (\$P2G02 * \$P2G03) since 2006: N/a

Variable Name	ISUBY\$\$
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. 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	$ISUBY\$ = (\$P2H02 * \$P2H03)$

Variable Name	IERET\$\$
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. 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	1984-1995: N/a 1996-2001: $IERET\$ = (\$P2I02 * \$P2I03)$ $(\$ = 96 - 01, \$ = M, N, \dots, R)$ since 2002: N/a

Variable Name	IMATY\$\$
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	<p>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.</p> <p>If the information about the number of months is missing, the sample mean of that variable has been assigned.</p> <p>In case of partial unit-non responding households this information has been imputed. For details see: Frick, Grabka & Groh-Samberg (2010).</p>
Format	<p>The value of this variable ranges from 0 to 999.999.</p> <p>This variable is in current year EURO.</p> <p>The original survey variables provided below can be found in the file _PKAL.</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	$IMATY\$ = (\$P2J02 * \$P2J03)$

Variable Name	ISTUY\$\$
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	<p>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.</p> <p>If the information about the number of months is missing, the sample mean of that variable has been assigned.</p> <p>In case of partial unit-non responding households this information has been imputed. For details see: Frick, Grabka & Groh-Samberg (2010).</p>
Format	<p>The value of this variable ranges from 0 to 999.999.</p> <p>This variable is in current year EURO.</p> <p>The original survey variables provided below can be found in the file _PKAL.</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	$\text{ISTUY}\$ = (\text{\$P2K02} * \text{\$P2K03})$

Variable Name	IMILT\$\$
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. 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	1984-1995: N/a since 1996: $IMILT\$ = (\$P2L02 * \$P2L03)$

Variable Name	IALIM\$\$
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: N/a since 2001: $IALIM\$ = (\$P2O02 * \$P2O03)$

Variable Name	IELSE\$\$
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. If the information about the number of months is missing, the sample mean of that variable has been assigned. In case of partial unit-non responding households this information has been imputed. For details see: Frick, Grabka & Groh-Samberg (2010).
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	$IELSE\$\$ = (\$P2M02 * \$P2M03)$

Variable Name	I13LY\$\$
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. 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	<p>I13LY\$\$ = Y13</p> <p>13th monthly salary variable list by survey year - each entry denoted in algorithm as Y13:</p> <p>ap3902 bp5902 cp5902 dp5902 ep5402 fp7202 gp7202 hp6702 ip6702 jp7702 kp7702 lp8202 mp6802 np6802 op5902 pp7702 qp7702 rp7702 sp7702 tp9502 up8002 vp10102 wp7802 xp9502 yp9602 zp9202</p>

Variable Name	I14LY\$\$
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	I14LY\$\$ = Y14 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 vp10104 wp7804 xp9504 yp9604 zp9204

Variable Name	IXMAS\$\$
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. 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	IXMAS\$\$ = YXMS 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 vp10106 wp7806 xp9506 yp9606 zp9206

Variable Name	IHOLY\$\$
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. 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	<p>IHOLY\$\$ = YHOL</p> <p>Vacation bonus variable list by survey year - each entry denoted in algorithm as YHOL:</p> <p>ap3908 bp5908 cp5908 dp5908 ep5408 fp7208 gp7208 hp6708 ip6708 jp7708 kp7708 lp8208 mp6808 np6808 op5908 pp7708 qp7708 rp7708 sp7708 tp9508 up8008 vp10108 wp7808 xp9508 yp9608 zp9208</p>

Variable Name	IGRAY\$\$
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. 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	IGRAY\$\$ = YGRY 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 vp10110 wp7810 xp9510 yp9610 zp9210

Variable Name	IOTHY\$\$
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: N/a since 1985: IOTHY\$\$ Other bonuses variable list by survey year: bp5912 cp5912 dp5912 ep5412 fp7212 gp7212 hp6712 ip6712 jp7712 kp7712 lp8212 mp6812 np6812 op5912 pp7712 qp7712 rp7712 sp7712 tp9512 up8012 vp10112 wp7812 xp9512 yp9612 zp9212

Variable Name	ITRAY\$\$
Variable Label	Commuting expenses, travel grant
Unit of Observation	Individual
Period	Annual
Description	This variable represents income from commuting expenses or travel grants provided by the employer of individuals in the household 16 years of age and older. In 1984-2005 specific questions related to this topic were not asked.
Method	Transcribed variable.
Format	The value of this variable ranges from 0 to 999.999. This variable is in current year EURO.
Algorithm	The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses. 1984-2005: N/a since 2006: ITRAY\$\$ Commuting or travel grants variable list by survey year: wp7902 xp9602 yp9702 zp9302

Variable Name	IDEMY\$\$
Variable Label	Indemnity payments
Unit of Observation	Individual
Period	Annual
Description	This variable represents income from indemnity payments of individuals in the household 16 years of age and older. In 1984-1990 specific questions related to this topic were not asked.
Method	Transcribed variable.
Format	The value of this variable ranges from 0 to 999.999. This variable is in current year EURO.
Algorithm	The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses. 1984-1990: N/a since 1991: IDEMY\$\$ Indemnity payments variable list by survey year: hp27g02 ip27g02 jp27g02 kp2402 lp2002 mp6502 np6502 op5602 pp7302 qp7302 rp7302 sp7302 tp9102 up7602 vp9702 wp8502 xp9102 yp9202 zp8802

Variable Name	IGRV1\$\$
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. 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: N/a since 1986: $IGRV1\$ = (\$P2D02 * YSPI)$ 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 vp10301 wp8001 xp9701 yp9801 zp9401

Variable Name	ISMP1\$\$
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. 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: N/a 1986-2001: $ISMP1\$ = (\$P2D02 * YSMP)$ since 2002: 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	ICIV1\$\$
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: N/a since 1986: $ICIV1\$ = (\$P2D02 * YCIV)$ 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 vp10303 wp8003 xp9703 yp9803 zp9403

Variable Name	IWAR1\$\$
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: N/a since 1986: $IWAR1\$ = (\$P2D02 * YWAR)$ 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 vp10305 wp8005 xp9705 yp9805 zp9413

Variable Name	IAGR1\$\$
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: N/a 1986-2001: $IAGR1\$ = (\$P2D02 * YAGR)$ since 2001: 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	IGUV1\$\$
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: N/a since 1986: $IGUV1\$ = (\$P2D02 * YGUV)$ 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 vp10307 wp8007 xp9707 yp9807 zp9411

Variable Name	IVBL1\$\$
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: N/a since 1986: $IVBL1\$ = (\$P2D02 * YVBL)$ 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 vp10309 wp8009 xp9709 yp9809 zp9405

Variable Name	ICOM1\$\$
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: N/a since 1986: $ICOM1\$ = (\$P2D02 * YCOM)$ 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 vp10311 wp8011 xp9711 yp9811 zp9407

Variable Name	IPRV1\$\$
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: N/a since 2003: $IPRV1\$ = (\$P2D02 * YPRV)$ Private pension variable list by survey year - each entry denoted in algorithm as YPRV: tp9713 up8213 vp10313 wp8013 xp9713 yp9813 zp9409

Variable Name	ISON1\$\$
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. 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: N/a since 1986: $ISON1\$ = (\$P2D02 * YSON)$ 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 vp10315 wp8015 xp9715 yp9815 zp9415

Variable Name	IGRV2\$\$
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: N/a since 1986: $IGRV2\$ = (\$P2E02 * YWID)$ 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 vp10302 wp8002 xp9702 yp9802 zp9402

Variable Name	ISMP2\$\$
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: N/a 1986-2001: $ISMP2\$ = (\$P2E02 * YSMP)$ since 2002: 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	ICIV2\$\$
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. 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: N/a since 1986: $ICIV2\$ = (\$P2E02 * YCIV)$ 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 vp10304 wp8004 xp9704 yp9804 zp9404

Variable Name	IWAR2\$\$
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: N/a since 1986: $IWAR2\$ = (\$P2E02 * YWAR)$ 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 vp10306 wp8006 xp9706 yp9806 zp9414

Variable Name	IAGR2\$\$
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: N/a 1986-2001: $IAGR2\$ = (\$P2E02 * YAGR)$ since 2001: 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	IGUV2\$\$
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: N/a since 1986: $IGUV2\$ = (\$P2E02 * YGUV)$ 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 vp10308 wp8008 xp9708 yp9808 zp9412

Variable Name	IVBL2\$\$
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. 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: N/a since 1986: $IVBL2\$ = (\$P2E02 * YVBL)$ 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 vp10310 wp8010 xp9710 yp9810 zp9406

Variable Name	ICOM2\$\$
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: N/a since 1986: $ICOM2\$ = (\$P2E02 * YCOM)$ 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 vp10312 wp8012 xp9712 yp9812 zp9408

Variable Name	IPRV2\$\$
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: N/a since 2003: $IPRV2\$ = (\$P2E02 * YPRV)$ Widows and orphans private pension variable list by survey year - each entry denoted in algorithm as YPRV: tp9714 up8214 vp10314 wp8014 xp9714 yp9814 zp9410

Variable Name	ISON2\$\$
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: N/a since 1986: $ISON2\$ = (\$P2E02 * YSON)$ 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 vp10316 wp8016 xp9716 yp9816 zp9416

Variable Name	FJOB1\$\$
Variable Label	Imputation flag: Wages, Salary from main job
Unit of Observation	Individual
Description	<p>This variable indicates if income from wages or salary from main job (IJOB1\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FJOB2\$\$
Variable Label	Imputation flag: Income from secondary job
Unit of Observation	Individual
Description	<p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FSELF\$\$
Variable Label	Imputation flag: Income from self-employment
Unit of Observation	Individual
Description	<p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FOLDY\$\$
Variable Label	Imputation flag: combined old-age, civil servants pensions.
Unit of Observation	Individual
Description	<p>This variable indicates if income from combined old-age, civil servants pensions (IOLDY\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FWIDY\$\$
Variable Label	Imputation flag: combined widows / orphans pension
Unit of Observation	Individual
Description	<p>This variable indicates if income from combined widows or orphans pension (IWIDY\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FCOMP\$\$
Variable Label	Imputation flag: Combined company pension
Unit of Observation	Individual
Description	<p>This variable indicates if income from combined company pension (ICOMP\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FPRVP\$\$
Variable Label	Imputation flag: Combined private pension
Unit of Observation	Individual
Description	<p>This variable indicates if income from combined private pension (IPRVP\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FUNBY\$\$
Variable Label	Imputation flag: Unemployment benefit
Unit of Observation	Individual
Description	<p>This variable indicates if income from unemployment benefit (IUNBY\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FUNAY\$\$
Variable Label	Imputation flag: Unemployment assistance
Unit of Observation	Individual
Description	<p>This variable indicates if income from unemployment assistance (IUNAY\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FSUBY\$\$
Variable Label	Imputation flag: Subsistence allowance
Unit of Observation	Individual
Description	<p>This variable indicates if income from subsistence allowance (ISUBY\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FERET\$\$
Variable Label	Imputation flag: Old-age transition benefit
Unit of Observation	Individual
Description	<p>This variable indicates if income from old-age transition benefit (IERET\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FMATY\$\$
Variable Label	Imputation flag: Maternity benefit
Unit of Observation	Individual
Description	<p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FSTUY\$\$
Variable Label	Imputation flag: Student grants
Unit of Observation	Individual
Description	<p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FMILT\$\$
Variable Label	Imputation flag: Military / community service pay
Unit of Observation	Individual
Description	<p>This variable indicates if income from military or community service pay (IMILT\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FALIM\$\$
Variable Label	Imputation flag: Alimony
Unit of Observation	Individual
Description	<p>This variable indicates if income from alimony (IALIM\$\$) 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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FELSE\$\$
Variable Label	Imputation flag: Private Transfers received
Unit of Observation	Individual
Description	<p>This variable indicates if income from private transfers received (IELSE\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	F13LY\$\$
Variable Label	Imputation flag: 13th monthly salary
Unit of Observation	Individual
Description	<p>This variable indicates if income from 13th monthly salary (I13LY\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	F14LY\$\$
Variable Label	Imputation flag: 14th monthly salary
Unit of Observation	Individual
Description	<p>This variable indicates if income from 14th monthly salary (I14LY\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FXMAS\$\$
Variable Label	Imputation flag: Christmas bonus
Unit of Observation	Individual
Description	<p>This variable indicates if income from Christmas bonus (IXMAS\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FHOLY\$\$
Variable Label	Imputation flag: Vacation bonus
Unit of Observation	Individual
Description	<p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FGRAY\$\$
Variable Label	Imputation flag: Profit-sharing / Gratifications
Unit of Observation	Individual
Description	<p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FOTHY\$\$
Variable Label	Imputation flag: Other bonuses
Unit of Observation	Individual
Description	<p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FTRAY\$\$
Variable Label	Imputation flag: Commuting expenses, travel grant
Unit of Observation	Individual
Description	<p>This variable indicates if income from commuting expenses or travel grants (ITRAY\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FDEMY\$\$
Variable Label	Imputation flag: Indemnity payments
Unit of Observation	Individual
Description	<p>This variable indicates if income from indemnity payments (IDEMY\$\$) has been imputed.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FGRV1\$\$
Variable Label	Imputation flag: statutory pension insurance
Unit of Observation	Individual
Description	<p>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.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	FGRV2\$\$
Variable Label	Imputation flag: widows or orphans statutory pension insurance
Unit of Observation	Individual
Description	<p>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.</p> <p>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).</p>
Method	<p>In the original SOEP data there are three types of missing values. These missing values can be interpreted as:</p> <ul style="list-style-type: none"> -1 = no answer or do not know -2 = does not apply -3 = original value was deleted because it was found to be implausible <p>The imputation procedures was used to fill in missing values represented by -1 (.A) and -3 (.C) only.</p>
Format	<p>0 = Not Imputed 1 = Fully Imputed</p> <p>This algorithm omits individuals with survey non-responses.</p>

Variable Name	M11101\$\$
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

```

array hosp{*} ap5101 bp7201 cp7201 dp7201 ep6901 fp8501 null hp8001 ip8001 null
               kp8701 lp9301 mp7901 np8301 op7201 pp100 qp99 rp99 sp98 tp101
               up94 vp107 wp98 xp101 yp111 zp98;
array nhosp{*} M11101$$;
do i = 1 to dim(netto);
  if netto{i} >= 10 & < 20 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	M11102\$\$
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	<pre> array hospd{*} ap5103 bpP7203 cp7203 dp7203 ep6903 fp8503 null hp8003 ip8003 null kp8703 lp9303 mp7903 np8303 op7203 pp10102 qp10002 rp10001 sp9901 tp10201 up9501 vp10801 wp9901 xp10201 yp11201 zp9901; array nhospd{*} M11102\$\$; do i = 1 to dim(netto); if netto{i} >= 10 & < 20 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; </pre>

Variable Name	M11103\$\$
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 several years the question related to this topic were not asked.
Method	Transcribed variable.
Format	<p>-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</p> <p>The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<pre> 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 null null null; array m03{*} M11103\$\$; do i = 1 to dim(netto); if netto{i} >= 10 & < 20 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; </pre>

Variable Name	M11104\$\$
Variable Label	Frequency play sports or exercise
Unit of Observation	Individual
Description	Indicates how often person plays sports, exercises or walks or swims. In several years the question related to this topic were not asked.
Method	Transcribed variable.
Format:	<p>-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</p> <p>The original survey variable provided below can be found in the _P files. This algorithm omits individuals with survey non-responses.</p>

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 vp0303 null xp0303 yp1815 zp0303;
array nsport{*} M11104$$
do i = 1 to dim(netto);
  if netto{i} >= 10 & < 20 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 (2,3,5,9,11,13,14,16,18,22,24,26) 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 (7,12,15,20,25) 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	M11105\$\$
Variable Label	Have had stroke
Unit of Observation	Individual
Description	Indicates whether a doctor ever diagnosed a stroke
Method	Transcribed variable
Format:	-2 = N/A - Child -1 = Item non-response 1 = has had a stroke
	<p>The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<p>1984-2008: Data not available in SOEP since 2009</p> <pre> array str{*} null array m05{*} m11105\$\$; do i = 1 to dim(netto); if netto{i} ge 10 & netto{i} le 19 then do; if str{i} gt 0 then m05{i}=str{i}; if str{i} in (-1,-3) then m05{i}=.M; if str{i} in (-2) then m05{i}=.S; end; else m05{i}=.S; end; </pre>

Variable Name	M11106\$\$
Variable Label	High blood pressure/circulation problems
Unit of Observation	Individual
Description	Indicates whether a doctor ever diagnosed a with high blood pressure or circulation problems
Method	Transcribed variable
Format:	<p>-2 = N/A - Child -1 = Item non-response 1 = Has or had problem with high blood pressure or circulation</p> <p>The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<p>1984-2008: Data not available in SOEP since 2009</p> <pre> array hbp{*} null null null null null null null null null null null null null null null null null null null null null null null null null zp10207; array m06{*} m11106\$\$; do i = 1 to dim(netto); if netto{i} ge 10 & netto{i} le 19 then do; if hbp{i} gt 0 then m06{i}=hbp{i}; if hbp{i} in (-1,-3) then m06{i}=.M; if hbp{i} in (-2) then m06{i}=.S; end; else m06{i}=.S; end; </pre>

Variable Name	M11107\$\$
Variable Label	Have or had diabetes
Unit of Observation	Individual
Description	Indicates whether a doctor ever diagnosed diabetes
Method	Transcribed variable
Format:	<p>-2 = N/A - Child</p> <p>-1 = Item non-response</p> <p>1 = Has or had problem with diabetes</p>
	<p>The original survey variables provided below can be found in the file _P.</p> <p>This algorithm omits individuals with survey non-responses.</p>
Algorithm	<p>1984-2008: Data not available in SOEP since 2009</p> <pre> array dia{*} null null null null null null null null null null null null null null null null null null null null null null null null null zp10201; array m07{*} m11107\$\$; do i = 1 to dim(netto); if netto{i} ge 10 & netto{i} le 19 then do; if dia{i} gt 0 then m07{i}=dia{i}; if dia{i} in (-1,-3) then m07{i}=.M; if dia{i} in (-2) then m07{i}=.S; end; else m07{i}=.S; end;</pre>

Variable Name	M11108\$\$
Variable Label	Have or had cancer
Unit of Observation	Individual
Description	Indicates whether a doctor ever diagnosed cancer
Method	Transcribed variable
Format:	<p>-2 = N/A - Child</p> <p>-1 = Item non-response</p> <p>1 = Has or had problem with cancer</p> <p>The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<p>1984-2008: Data not available in SOEP since 2009</p> <pre> array can{*} null null null null null null null null null null null null null null null null null null null null null null null null null null null null null null null array m08{*} m11108\$\$; do i = 1 to dim(netto); if netto{i} ge 10 & netto{i} le 19 then do; if can{i} gt 0 then m08{i}=can{i}; if can{i} in (-1,-3) then m08{i}=.M; if can{i} in (-2) then m08{i}=.S; end; else m08{i}=.S; end;</pre>

Variable Name	M11109\$\$
Variable Label	Have or had psychiatric problems
Unit of Observation	Individual
Description	Indicates whether a doctor ever diagnosed a depressive diseases.
Method	Transcribed variable
Format:	<p>-2 = N/A - Child</p> <p>-1 = Item non-response</p> <p>1 = Has or had problem with depressive diseases</p> <p>The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<p>1984-2008: Data not available in SOEP since 2009</p> <pre> array psy{*} null null null null null null null null null null null null null null null null null null null null null null null null null null null null null null array m09{*} m11109\$\$; do i = 1 to dim(netto); if netto{i} ge 10 & netto{i} le 19 then do; if psy{i} gt 0 then m09{i}=psy{i}; if psy{i} in (-1,-3) then m09{i}=.M; if psy{i} in (-2) then m09{i}=.S; end; else m09{i}=.S; end;</pre>

Variable Name	M11110\$\$
Variable Label	Have or had arthritis
Unit of Observation	Individual
Description	Indicates whether person has or had problems with arthritis
Method	n.a.
Format:	Information is <u>not</u> available in the SOEP
Algorithm	n.a.

Variable Name	M11111\$\$
Variable Label	Angina or heart condition
Unit of Observation	Individual
Description	Indicates whether a doctor ever diagnosed angina or heart condition problems
Method	Transcribed variable
Format:	<p>-2 = N/A - Child</p> <p>-1 = Item non-response</p> <p>1 = Has or had problem with depressive diseases</p> <p>The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<p>1984-2008: Data not available in SOEP since 2009</p> <pre> array ang{*} null null null null null null null null null null null null null null null null null null null null null null null null null null null null null null array m11{*} m11111\$\$; do i = 1 to dim(netto); if netto{i} ge 10 & netto{i} le 19 then do; if ang{i} gt 0 then m11{i}=ang{i}; if ang{i} in (-1,-3) then m11{i}=.M; if ang{i} in (-2) then m11{i}=.S; end; else m11{i}=.S; end;</pre>

Variable Name	M11112\$\$
Variable Label	Have or had asthma or breathing difficulty
Unit of Observation	Individual
Description	Indicates whether person has or had problems with asthma or breathing difficulties
Method	n.a
Format:	n.a.
Algorithm	Information is <u>not</u> available in the SOEP

Variable Name	M11113\$\$
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 several years the question 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

```

null=-2;
*First set missing values for all other data;
M1111384 - M1111301, M1111303, M1111305, M1111307= -2;

array hlpstair{*} sp87 up84 wp88 yp100;
array hlthaffstair {*} M1111302 M1111304 M1111306 M1111308;
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	M11114\$\$
Variable Label	Have difficulty or need help of others to bathe
Unit of Observation	Individual
Description	Indicates whether person has trouble with or needs help of others to bathe
Method	n.a
Format:	n.a.
Algorithm	Information is <u>not</u> available in the SOEP

Variable Name	M11115\$\$
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:	<p>-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</p> <p>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.</p>
Algorithm	<p>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."</p> <pre> if erhebj=&year; 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{*} simp\$\$; array M15{*} M11115\$\$; do i = 1 to dim(netto); if netto{i} >= 10 & < 20 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; </pre>

Variable Name	M11116\$\$
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:	<p>-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</p> <p>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.</p>
Algorithm	<pre> 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, ...) 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{*} bed\$\$; array M16{*} M11116\$\$; do i = 1 to dim(netto); if netto{i} >= 10 & < 20 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; </pre>

Variable Name	M11117\$\$
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 "mult-grad" and "maxgrad" which can be found in the file PFLEGE. This algorithm omits individuals with survey non-responses.

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, ...)
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{*} outh$$
array M17{*} M11117$$;
do i = 1 to dim(netto);
if netto{i} >= 10 & < 20 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	M11118\$\$
Variable Label	Walk 10+ minutes difficult
Unit of Observation	Individual
Description	Indicates whether person has trouble to walk unaided for 10 or more minutes
Method	n.a
Format:	n.a.
Algorithm	Information is <u>not</u> available in the SOEP

Variable Name M11119\$\$

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 "multgrad" and "maxgrad" which can be found in the file PFLEGE.
This algorithm omits individuals with survey non-responses.

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, ...)
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{*} inh$$;
array M19{*} M11119$$;
do i = 1 to dim(netto);
if netto{i} >= 10 & < 20 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	M11120\$\$
Variable Label	Health limits bending, lifting, stooping
Unit of Observation	Individual
Description	Indicates whether a person's health limits his ability to bend, lift, or stoop
Method	n.a
Format:	n.a.
Algorithm	Information is <u>not</u> available in the SOEP

Variable Name	M11121\$\$
Variable Label	Health limits vigorous physical activity
Unit of Observation	Individual
Description	Indicates whether person's health limits vigorous physical activity
Method	n.a
Format:	n.a.
Algorithm	Information is <u>not</u> available in the SOEP

Variable Name	M11122\$\$
Variable Label	Height in centimeters
Unit of Observation	Individual
Description	This variable indicates each person's height in centimeters on the interview date. In 1984-2001, 2003/05/07/09 questions related to this topic were not asked.
Method	Transcribed variable.
Format	-2 = N/A - Child -1 = Item non-response 0 to 220 = Height in centimeters The original survey variables provided below can be found in the file _P. This algorithm omits individuals with
Algorithm	1984-2001, 2003, 2005, 2007, 2009: Data not available in SOEP 2002: M1112202=sp90 2004: M1112204=up90 2006: M1112206=wp94 2008: M1112208=yp107

Variable Name	M11123\$\$
Variable Label	Weight in kilos
Unit of Observation	Individual
Description	This variable indicates each person's weight in kilos on the interview date. In 1984-2001, 2003/05/07/09 questions related to this topic were not asked.
Method	Transcribed variable.
Format	-2 = N/A - Child -1 = Item non-response 0 to 400 = Weight in kilos
	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, 2005, 2007, 2009: Data not available in SOEP 2002: M1112302=sp91 2004: M1112304=up91 2006: M1112306=wp95 2008: M1112308=yp108

Variable Name	M11124\$\$
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 vp10502 wp9602
xp10502 yp10902 zp9602

*This step is to assign a value for 1990 using 1991 degree of disability;
if gnetto ge 10 & < 30 & ghnetto=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 ge 10 & < 30 & jhnetto=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 >= 10 & < 30 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 ge 30 then M11124\$\$=1;
else M11124\$\$=0;

Variable Name	M11125\$\$
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.
Format	<p>-2 = N/A - Child -1 = Item non-response 0 = totally unhappy,, 10 = totally happy</p> <p>The original survey variables provided below can be found in the file __P. This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions: D11101__ = Age of Individual</p>

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
vp0101 wp0101 xp0101 yp0101 zp0101;

```

if D11101$$ ge 16 then do                                (Y=a-x, $$=84, ...)
  if Ysatis ge 0 then M11125$$ = Ysatis
  else                M11125$$ = .M
end
if D11101$$ lt 16 then M11125$$ = .C

```

Variable Name	M11126\$\$
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:	<p>-2 = N/A - Child -1 = Item non-response 1 = Excellent 2 = Very good 3 = Good 4 = Fair 5 = Poor</p> <p>The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<p>1984-1991, 1993: Data not available in SOEP 1992, since 1994:</p> <pre> array sta1{*} null null null null null null null ip77 null kp83 lp89 mp75 np79 op66 pp95 qp95 rp95 sp86 tp98 up83 vp104 wp87 xp104 yp99 zp95; array stat{*} M11126\$\$ do i = 1 to dim(netto); if netto{i} >= 10 & < 20 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; </pre>

Variable Name	M11127\$\$
Variable Label	Number of doctor visits in previous year
Unit of Observation	Individual
Description	<p>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.</p> <p>In 1990 and 1993 the Number of doctor visits were not asked.</p>
Method	Transcribed variable
Format:	<p>-2 = N/A - Child</p> <p>-1 = Item non-response</p> <p>0 to 400 = doctors visits in previous year.</p> <p>The original survey variables provided below can be found in the file _P. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<p>1990, 1993: Data not available in SOEP</p> <p>1984-1989, 1991-1992, since 1994:</p> <pre> 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 vp10601 wp9701 xp10601 yp11001 zp9701; array s101{*} M11127\$\$ do i = 1 to dim(netto); if netto{i} >= 10 & < 20 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; </pre>

Variable Name	P11101\$\$
Variable Label	Overall life satisfaction
Unit of Observation	Individual
Description	This variable indicates the satisfaction with life in general 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 life in general on a scale of 0 to 10, in which 0 means completely dissatisfied and 10 means completely satisfied.
Format	<p>-2 = N/A - Child -1 = Item non-response 0 = completely dissatisfied,, 10 = completely satisfied</p> <p>The original survey variables provided below can be found in the file __P. This algorithm omits individuals with survey non-responses.</p> <p>Equivalent Data File Variable Definitions: D11101__ = Age of Individual</p>
Algorithm	<p>Satisfaction With Life variable list by survey year - each entry denoted in algorithm as Ysatis:</p> <p>ap6801 bp9301 cp9601 dp9801 ep89 fp108 gp109 hp10901 ip10901 jp10901 kp10401 lp10401 mp11001 np11701 op12301 pp13501 qp14301 rp13501 sp13501 tp14201 up14501 vp154 wp142 xp149 yp15501 zp15701</p> <pre> if D11101\$\$ ge 16 then do (Y=a-x, \$\$=84, ...) if Ysatis ge 0 then P11101\$\$ = Ysatis else Ü11101\$\$ = .M end if D11101\$\$ lt 16 then P11101\$\$ = .C </pre>

Variable Name	W11101\$\$
Variable Label	Individual Weight – Respondent Individuals
Unit of Observation	Individual
Description	Indicates the individual's population and sample weight. The individual weight provides a time-series consistent information, which might be used for analyses of income and income inequality over time. Given the selectivity built in the answering behavior of first-time respondents with respect to income questions (showing a significant higher share of item-non response which is especially true for a first wave of a new sub-sample), this weighting factor excludes every first wave of a new sub-sample of the SOEP. If all sub-samples will be used, the appropriate variable is W11105\$\$.
Method	Individual weights to compensate for unequal probabilities of selection and sample attrition are necessary to obtain populations based statistics. The individual weights also encompass population weights.
Format	<p>The value of this variable ranges from 0 to 80.000.</p> <p>The original survey variables provided below can be found in the file PHRF. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<p>W11101\$\$ = Yphrf</p> <p> 1995 W11101\$\$ = Lphrf * 1.056 if X11104LL in (21,22,23), else=0 1998 W11101\$\$ = Ophrfad 2000 W11101\$\$ = Qphrfae 2002 W11101\$\$ = Sphrfaf 2006 W11101\$\$ = Wphrfag </p>

Variable Name	W11102\$\$
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	<p>The value of this variable ranges from 0 to 80.000.</p> <p>The original survey variables provided below can be found in the file HHRF. This algorithm omits individuals with survey non-responses.</p>
Algorithm	$W11102\$ = Yhhrf$

Variable Name	W11103\$\$
Variable Label	Longitudinal Weight – Respondent Individuals
Unit of Observation	Individual
Description	Indicates the individual's longitudinal sample weight.
Method	<p>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.</p> <p>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.</p> <p>Example: longitudinal weight = W1110186*W1110387*W1110388*W1110389*W1110390</p> <p>The created longitudinal weight should be used with any longitudinal sample.</p>
Format	<p>N/A</p> <p>The original survey variables provided below can be found in the file PHRF. This algorithm omits individuals with survey non-responses.</p>
Algorithm	<p>1984: N/A</p> <p>since 1985: W11103\$\$ = Ypbleib</p>

Variable Name	W11104\$\$
Variable Label	Population Factor
Unit of Observation	Household
Description	This correction factor is already integrated into W11101\$\$ and is not repeated here.
Format	Data <u>not</u> available in SOEP

Variable Name	W11105\$\$
Variable Label	Individual Weight – all samples
Unit of Observation	Individual
Description	Indicates the individual's weight for all sub-samples of the SOEP.
Method	<p>Individual weights to compensate for unequal probabilities of selection and sample attrition are necessary to obtain populations based statistics.</p> <p>These weights should be used when analyses include all sub-samples of the SOEP, namely samples A thru H.</p> <p>For a more detailed discussion of weighting issues in the SOEP please see <i>Desktop Companion to the German Socio-Economic Panel Study (SOEP)</i>.</p>
Format	<p>The value of this variable ranges from 0 to 80.000.</p> <p>The original survey variables provided below can be found in the file PHRF. This algorithm omits individuals with survey non-responses.</p>
Algorithm	W11105\$\$ = \$PHRF

Variable Name	Variable Label
W11107\$\$	Cross-sectional Weight – Enumerated Individuals
W11108\$\$	Longitudinal Weight – Enumerated Individuals
W11109\$\$	Population Factor for W11103\$\$
W11110\$\$	Population Factor for W11107\$\$
W11111\$\$	Population Factor for W11108\$\$
Algorithm	Data <u>not</u> available in SOEP

References

- Hagenaars A.J.M., K. de Vos, M. A. 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., Grabka, Markus M. und Groh-Samberg, Olaf (2010): *Dealing with Incomplete Household Panel Data in Inequality Research*. *SOEP Papers on Multidisciplinary Panel Data Research at DIW Berlin*, No. 290, Berlin (see: http://www.diw.de/documents/publikationen/73/diw_01.c.354683.de/diw_sp0290.pdf)
- 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.
- 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.