

DIW Berlin / SOEP (Ed.)

**Research Report**

## SOEP 2014 - Documentation of the pooled dataset on children in KIDLONG for SOEP v31.1

SOEP Survey Papers, No. 303

**Provided in Cooperation with:**

German Institute for Economic Research (DIW Berlin)

*Suggested Citation:* DIW Berlin / SOEP (Ed.) (2016) : SOEP 2014 - Documentation of the pooled dataset on children in KIDLONG for SOEP v31.1, SOEP Survey Papers, No. 303, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<http://hdl.handle.net/10419/152307>

**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.*



<http://creativecommons.org/licenses/by-sa/4.0/>

## SOEP Survey Papers

Series D – Variable Descriptions and Coding

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

2016

# SOEP 2014 – Documentation of the Pooled Dataset on Children in KIDLONG for SOEP v31.1

SOEP Group

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

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

**Series A** – Survey Instruments (Erhebungsinstrumente)

**Series B** – Survey Reports (Methodenberichte)

**Series C** – Data Documentation (Datendokumentationen)

**Series D** – Variable Descriptions and Coding

**Series E** – SOEPmonitors

**Series F** – SOEP Newsletters

**Series G** – General Issues and Teaching Materials

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

#### **Editors:**

Dr. Jan Goebel, DIW Berlin

Prof. Dr. Martin Kroh, DIW Berlin and Humboldt Universität Berlin

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

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

Please cite this paper as follows:

SOEP Group. 2016. SOEP 2014 – Documentation of the Pooled Dataset on Children in KIDLONG for SOEP v31.1. SOEP Survey Papers 303: Series D. Berlin: DIW/SOEP



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

© 2016 by SOEP

ISSN: 2193-5580 (online)

DIW Berlin

German Socio-Economic Panel (SOEP)

Mohrenstr. 58

10117 Berlin

Germany

soeppapers@diw.de

# SOEP 2014 – Documentation of the Pooled Dataset on Children in KIDLONG for SOEP v31.1

**SOEP Group**

*2016*

The file KIDLONG is part of a collection, which is released with doi: [10.5684/soep.v31.1](https://doi.org/10.5684/soep.v31.1)

## Contents

<b>1 Overview</b>	<b>3</b>
<b>2 Identifier</b>	<b>3</b>
syear – Survey year . . . . .	3
hhnr – Original household number . . . . .	3
hhnrakt – Wave-specific household id . . . . .	3
persnr – Id of child . . . . .	3
<b>3 Basic demographics (var names starting with K)</b>	<b>3</b>
k_samp – Sample membership . . . . .	3
k_nat – Nationality . . . . .	5
k_inhh – Member of household . . . . .	5
k_size – Household size . . . . .	6
k_nrkid – Number of children in household . . . . .	6
k_rel – Relationship to household head . . . . .	6
k_nrinhh – Child id within household . . . . .	8
<b>4 Pointers (var names starting with K)</b>	<b>9</b>
k_phead – Pointer to household head . . . . .	9
k_pheadp – Pointer to partner of household head . . . . .	9
k_pmum – Pointer to mother . . . . .	9
k_pmump – Pointer to partner of mother . . . . .	10
k_instr – Language version of questionnaire . . . . .	10
<b>5 Variables on institutional arrangements (var names starting with KS)</b>	<b>10</b>
ks_pre – Pre-school . . . . .	10
ks_gen – General school . . . . .	11
ks_vocat – Vocational school . . . . .	11
ks_none – None of these institutions . . . . .	12
ks_east – General school east germany . . . . .	12
ks_eastp – Pre-school east germany . . . . .	12
ks_other – Other type of school . . . . .	13
<b>6 Variables on day care arrangements (var names starting with KD)</b>	<b>13</b>
kd_publ – Provider: public facility . . . . .	13
kd_priv – Provider: church facility . . . . .	14
kd_indep – Provider: private non-profit facility . . . . .	14
kd_comp – Provider: company facility . . . . .	15
kd_comm – Provider: private-commercial facility . . . . .	15
kd_rely – Adherence to opening time . . . . .	15
kd_time – Time spent at facility . . . . .	16
kd_hrs – Hours p.day spent at facility . . . . .	16
kd_lunch – Lunch at facility . . . . .	16
<b>7 Variables on cost of school (var names starting with KS)</b>	<b>16</b>
ks_lunch – Lunch at school . . . . .	17
ks_amtp – Costs for school per month . . . . .	17
ks_cost – School: no costs . . . . .	17

<b>8 Variables on costs of pre-school (var names starting with KS)</b>	<b>17</b>
kk_amtp – Costs for pre-school per month . . . . .	17
kk_cost – Pre-school: no costs . . . . .	17
<b>9 Variables on cost of day care arrangements (var names starting with KD)</b>	<b>18</b>
kd_fee – Payment of fee . . . . .	18
kd_amtp – Amount paid . . . . .	18
<b>10 Variables describing carer outside the household (var names starting with KC)</b>	<b>18</b>
kc_relat – Carer: relatives . . . . .	18
kc_frnd – Carer: friends . . . . .	19
kc_paid – Carer: paid . . . . .	19
kc_mindr – Carer: child minder . . . . .	19
kc_none – Carer: none . . . . .	20
<b>11 Activities below the age of 6 (var names starting with KA06)</b>	<b>20</b>
ka06_spo – Gymnastics under 6 . . . . .	20
ka06_mus – Music education under 6 . . . . .	20
ka06_art – Art class under 6 . . . . .	20
ka06_oth – Other activities under 6 . . . . .	21
ka06_non – No activities under 6 . . . . .	21
<b>12 Activities for children aged 6 to 16 (var names starting with KA16)</b>	<b>21</b>
ka16_ssp – School sports club above 6 . . . . .	21
ka16_smu – School music club above 6 . . . . .	21
ka16_sar – School art club above 6 . . . . .	22
ka16_sth – School theater club above 6 . . . . .	22
ka16_sot – Other school activities above 6 . . . . .	22
ka16_spo – Sports club above 6 . . . . .	22
ka16_mus – Music education above 6 . . . . .	22
ka16_art – Art class above 6 . . . . .	23
ka16_org – Volunteer fire dept, red cross etc. Above 6 . . . . .	23
ka16_yth – Youth group above 6 . . . . .	23
ka16_ctr – Youth center above 6 . . . . .	23
ka16_non – No activities above 6 . . . . .	23
<b>13 Variables on costs of school, carer, activities etc (var names starting with K)</b>	<b>24</b>
k_amtp – Costs per month . . . . .	24
k_cost – No costs . . . . .	24
<b>14 Variables on costs of external care (var names starting with KC)</b>	<b>24</b>
kc_amtp – External care: costs per month . . . . .	24
kc_cost – External care: no costs . . . . .	25

## 1 Overview

The variables stored in the **KIDLONG** file are based on the information annually collected and stored in the wave-specific **\$KIND** files. The relevant information is not provided by children themselves but by answers to the questions in the household questionnaire given by the respondent within the household (mostly the head of the household). This data is reaggregated at the person level and stored as child-specific entries in the file **\$KIND**. In order to create variables that comprise consistent longitudinal information the variables from the annual **\$KIND** files need to be modified, meaning the information needs to be adjusted to be comparable over time. Since the annual **\$KIND** datasets also contain additional information on institutional care and school attendance for children and young people; in selected waves, questions on the costs of childcare as well as non-institutional care arrangements are also asked. The entries in **KIDLONG** are identifiable with the codes “20 to 25” in the wave-specific variable **\$NETTO** (in the file **PPFAD**) or the code “2” in the variable **\$NETOLD**. Please note that “year of birth”, “month of birth” and “sex” are not included in this dataset. A child’s year of birth (**GEBJAHR**), month of birth (**GEMONAT**) and child’s sex (**SEX**) are included in **PPFAD**. In **PPFAD** you will find consistent longitudinal information. **KIDLONG** comprises data from waves **A .....BE**, i.e. survey years 1984-2014. **KIDLONG** is a pooled dataset with the identifiers **persnr** and **syear**.

## 2 Identifier

**syear** – Survey year

---

Variable **syear** depicts the year of observation. It ranges from 1984 to the most recent wave of data (currently 2014). The variable **syear** is needed in conjunction with **persnr** unique identification of a PERSON-YEAR observation in **KIDLONG**.

**hhnr** – Original household number

---

**hhnrakt** – Wave-specific household id

---

**persnr** – Id of child

---

Unique person identifier for each person in the SOEP database.

## 3 Basic demographics (var names starting with K)

**k\_samp** – Sample membership

---

1	A German HH head, West	44942
2	B1 Turkish HH head	10518
3	B2 Yugoslavian HH head	3518
4	B3 Greek HH head	2228
5	B4 Italian HH head	4214
6	B5 Spanish HH head	1689
7	C German HH head, East	20795
8	D1 Immigrant W. Germany Since 1984	2699
9	D2 Immigrant W. Germany Since 1984	2763
10	E1, PAPI	1960
11	E2, CAPI	1916

12	F2 ISOEP, Addtnl Foreigner Enumeration	20438
13	F1 ISOEP, Innovation	1370
14	G1 High Income (PreTest)	0
15	G2 High Income	3592
...	(5 rows omitted)	9201
23	L Special Survey 2012	0
24	M1 Migration 2013	4260
60	L2/ 3 FiD Screening 2010/ II Low Income	6374
61	L2/ 3 FiD Screening 2010/ II Single Parent	4616
62	L2/ 3 FiD Screening 2010/ II Families with Several Children	11406
63	L1 FiD Birth Cohort 2007	4103
64	L1 FiD Birth Cohort 2008	4089
65	L1 FiD Birth Cohort 2009	3586
66	L1 FiD Birth Cohort 2010	3666
-1	No Answer	0
-2	Does not apply	0
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **k\_samp** depicts the sample membership of each child. It contains unchanging sample information: for example, a person from sample C would retain this original sample number (the sample in which the household first entered SOEP) even after having later moved to West Germany (see also the comments on the variables **ssampreg** in the documentation of PPFAD). Similarly, a person from sample B who acquired a new nationality would still remain in his or her original sample.

Value	label
1	A German HH head, West
2	B1 Turkish HH head
3	B2 Yugoslavian HH head
4	B3 Greek HH head
5	B4 Italian HH head
6	B5 Spanish HH head
7	C German HH head, East
8	D1 Immigrant W. Germany Since 1984
9	D2 Immigrant W. Germany Since 1984
10	E1, PAPI
11	E2, CAPI
12	F1 ISOEP, Innovation
13	F2 ISOEP, Addtnl Foreigner Enumeration
14	G1 High Income (PreTest)
15	G2 High Income
16	H Refreshment Sample 2006
19	I Incentivation 2009
20	J Increase 2011
21	K Increase 2012
22	I2 Increase IS 2012
23	L Special Survey 2012
24	M1 Migration 2013



Value	label
60	L2/ 3 FiD Screening 2010/ II Low Income
61	L2/ 3 FiD Screening 2010/ II Single Parent
62	L2/ 3 FiD Screening 2010/ II Families with Several Children
63	L1 FiD Birth Cohort 2007
64	L1 FiD Birth Cohort 2008
65	L1 FiD Birth Cohort 2009
66	L1 FiD Birth Cohort 2010

### k\_nat – Nationality

0	Nationality of household head	957
1	German	242
2	Turkish	653
3	Yugoslavian	262
4	Greek	177
5	Italian	289
6	Spanish	147
11	French	2
12	Belgian, Dutch, Luxembourg	1
15	Swedish	3
18	American	2
23	Korean	3
24	Iranian	1
25	Indonesian	2
28	Portugese	1
29	Bulgarian	3
39	Israelian	6
43	Afghan	2
-1	No answer	83
-2	Does not apply	171107
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

Please note that this variable was only included in the survey in wave A (1984) and wave E (1988).

### k\_inhh – Member of household

0	Lives in household	148666
2	Not in hh due to training, college	11
3	Not in hh due to job	3
4	Not in hh due to hospital stay	1
5	Absent for longer	14
6	Not In hh due to other reasons	19
11	Born	6826

12	Moved within Germany since last survey	724
13	Moved f.foreign country since l.survey	127
14	Returned to hh	168
15	Moved within Germany prior to l.survey	18
16	Moved f.foreign country prior l.survey	16
17	Born before last survey	124
18	Moved from former GDR	134
20	Lives in new hh	835
30	Lives in hh for a long time	15498
31	Born last year	392
32	Moved from within Germany last year	152
33	Moved from foreign country last year	13
38	Moved f.former GDR since jan. prev. year	38
40	Now living In hh	112
-1	No answer	38
-2	Does not apply	13
-3	Answer improbable	1
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

### **k\_size** – Household size

---

The variable **k\_size** is a copy of the variable **shhgr** from the wave-specific files \$HBRUTTO.

### **k\_nrkid** – Number of children in household

---

For each child in **KIDLONG** the variable **k\_nrkid** provides the total number of children up to the age 16 in the current wave's household.

### **k\_rel** – Relationship to household head

---

0	Head of household	0
1	Spouse of hh head	1
2	Partner of hh head	1
3	Son, daughter of hh head	132022
4	Foster child of hh head	424
5	Son, daughter-in-law of hh head	3
8	Brother, sister of hh head	26
9	Grandchild of hh head	2177
10	Other relative of hh head	129
11	Other non-relative of hh head	1154
12	Child of hh head partner	2084
13	Same-sex spouse	0
21	Son, daughter	34590
22	Stepchild (child of the partner)	749
23	Adoptive child	46
...	(14 rows omitted)	497
45	Foster brother/ sister	1
51	Brother, sister-in-law (spouse of brother/ sister)	0
52	Brother, sister -in-law (brother/ sister of spouse)	1

61	Aunt, uncle	0
62	Niece/ nephew	18
63	Cousin/ cousine	1
64	Other relative	1
71	Others	6
99	Unknown	0
-1	No answer	7
-2	Does not apply	5
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **k\_rel** indicates the wave-specific relationship of the child to head of household and can be found as **\$stell** in file **\$PBRUTTO**. The values (1) and (2) should by definition not be relevant in the file **KIDLONG**, yet two observations in the early periods (1984 and 1987) are valid values “spouse” or “life partner” of the head of household. Please note that in 2012 **bcstell** in **BCPBRUTTO** changed. Thus for 2012 the values 21 - 62 indicate the wave-specific relationship of the child to the head of the household.  
\$stell in \$PBRUTTO

value	label
0	Head of HH
11	Husband / wife
12	Same-sex partner
13	Life partner
21	Biological child
22	Stepchild (child of the partner)
23	Adoptedchild
24	Fosterchild
25	Grandchild
26	Great-grandchild
27	Son, daughter-in-law
31	Biological mother / father
32	Step father / step mother / spouse of father or mother
33	Adoptive father or mother
34	Fosterfather or mother
35	Parent-in-law
36	Grandfather, -mother
41	Biological brother / sister
42	Half-brother, half-sister
43	Stepbrother, stepsister
44	Adoptive brother / sister
45	Fosterbrother / sister
51	Brother, sister -in-law (spouse of brother / sister)
52	Brother, sister -in-law (brother / sister of spouse)
61	Aunt / uncle
62	Niece / nephew
63	Cousin / cousine
64	Other relatives
71	Other

value	label
99	Unknown

\$k\_rel in KIDLONG valid from 2012

value	label
0	Head of household
1	Spouse of hh head
2	Partner of hh head
3	Son, daughter of hh head
4	Foster child of hh head
5	Son, daughter-in-law of hh head
8	Brother, sister of hh head
9	Grandchild of hh head
10	Other relative of hh head
11	Other non-relative of hh head
12	Child of hh head partner
13	Partner
21	Son, daughter
22	Stepchild (child of the partner)
23	Adoptive child
24	Foster child
25	Grandchild
26	Great-grandchild
27	Son, daughter-in-law
31	Biological mother / father
32	Stepfather / mother or partner of mother / father
33	Adoptive father or mother
34	Foster father or mother
35	Parent-in-law
36	Grandfather, -mother
41	Brother, sister
42	Half-brother, half-sister
43	Stepbrother, stepsister
44	Adoptive brother / sister
45	Foster brother / sister
51	Brother, sister-in-law (spouse of brother / sister)
52	Brother, sister -in-law (brother / sister of spouse)
61	Aunt, uncle
62	Niece / nephew
63	Cousin / cousin
64	Other relative
71	Others
99	Unknown

### k\_nrinh – Child id within household

The variable **k\_inhh** indicates the wave-specific household affiliation of a child and can be found as **spzug** in the file **\$PBRUTTO**. The values (19) and (39) have been recoded in

the file **KIDLONG**, and are now code “-1” (No answer).

## 4 Pointers (var names starting with K)

The pointers provided in **KIDLONG** comprise **persnr** of the household head of the child **\$khv**, the partner of the household head **\$khvp**, the mother of the child **\$kmutti**, and the partner of the mother **\$kmup**. The first two pointers **\$khv** and **\$khvp** are using the codes (0) and (1)/(2) of **\$STELL** in the wave-specific **\$PBRUTTO**. After extensive testing of the distribution of the data up to that point, the pointers to mother (variable **\$kmutti**) and mother’s partner (variable **\$kmup**) were revised in July 1994 for SOEP waves A (1984) to J(1993), since in some cases these variables showed a very high percentage of missings. The identifiers linking the **PERSNR** of a child’s mother to each child in **bekind**, are based on the variables **\$stell**, **partnr\$\$** in the wave-specific files **\$PGEN** and if possible information from the files **BIOBIRTH** or **BIOAGE01**. More information on **bekmutti** is included in the comment section of **\$kmutti**.

### **k\_phead** – Pointer to household head

---

Generated as **persnr** of the head of household, i.e. person with **\$stell** = 0 in the wave-specific file **\$PBRUTTO**.

### **k\_pheadp** – Pointer to partner of household head

---

The variable **persnr** of the partner of the head of household combines information of **\$stell** in the wave-specific file **\$PBRUTTO** with where applicable the partner information, i.e. pointer to spouse or partner: variable **PARTNR\$\$** in the wave-specific file **\$PGEN**. The variable includes -3 as a missing value for the years 1985/1986/1988/1989 in **k\_samp A** and **B**.

### **k\_pmum** – Pointer to mother

---

The variable **persnr** of a child’s mother. After extensive testing of the distribution of the data up to that point, the pointers to mother (variable **\$kmutti**) and mother’s partner (variable **\$kmup**) were revised in July 1994 for SOEP waves A (1984) to J (1993), since in some cases these variables showed a very high percentage of missings. The procedure for generating the pointers was modified to define the mother-child relationship as a combination of the variables **\$stell** (=relationship to household head) and **PARTNR\$\$** (=person ID number of spouse or partner). The information of **PARTNR\$\$** in the wave-specific file **\$PGEN** is only available when an interview took place with the partner. If this data is lacking the combination of relationship to household head of the mother and of the mother’s partner is used. In this context, for waves A to J, the newly generated pointers were tested longitudinally over time for consistency. While it is quite realistic that a mother would change partners over time, (e.g., second marriage after partner’s death, or separation from partner), it seems less plausible that there would be pointers to different mothers for the different survey years. According to a strictly biological definition of motherhood, this would certainly not be possible, but according to a social definition of motherhood, it would be possible to have a new “social” mother after the biological mother dies or moves out. This problem is quantitatively insignificant, however (e.g., a changing mother pointer

was only identified for 29 children from wave A to J). Further information on the definition of social/ biological mother-child relationships in the SOEP can be found in the BIO documentation for the file BIOBIRTH.

### **k\_pmump** – Pointer to partner of mother

---

The variable **persnr** of the partner of the child's mother (cf. PARTNR\$\$ of the mother in the person-related file \$PGEN).

### **k\_instr** – Language version of questionnaire

---

1	German	13928
2	Turkish	3454
3	Yugoslav	1239
4	Greek	801
5	Italian	1352
6	Spanish	571
7	German (East)	6439
8	Immigrant	2356
-1	No answer	0
-2	Does not apply	143803
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

## **5 Variables on institutional arrangements (var names starting with KS)**

### **ks\_pre** – Pre-school

---

1	Yes, Kindergarten, nursery	38312
2	Part-time	1279
3	In the morning	1630
4	In the afternoon	270
5	Full-time	2709
-1	No answer	1870
-2	Does not apply	127873
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

In order to integrate various information on formal childcare from all waves (A ...BE) into one variable **ks\_pre**, the different codes of different questionnaire items were modified. For the years 1984, 1985, 1986 and 1995-2014 value 1 gives the answer “yes, Kindergarten, creche, nursery”. For waves D ...G formal childcare was surveyed as part time or full time attendance. Since full time attendance was also surveyed in waves H ...K the value 2 was recoded to 5. This allows to code in the morning and in the afternoon attendance as value 3 and 4 for waves H ...K. For the remaining waves L ...Z the answers were coded accordingly. Please note that from wave L to P the question regarding formal care did distinguish

between kindergarten, creche, nursery and child minder. The children who were looked after by a child minded are coded -2 in **ks\_pre** and can be found in variable **kc\_mindr** (see below section “Variables on day care arrangements”). Additionally, in wave Z value 2 in the wave-specific file ZKIND was used for the answer no, not attending. Children who do not attend formal childcare in wave Z are coded -2 in variable **ks\_pre** and can be found in variable **ks\_none**. This is true for all subsequent waves following Z.

### **ks\_gen** – General school

---

1	Primary school	39251
2	Lower secondary school	12488
3	Secondary school	15287
4	Upper secondary school	17521
5	Comprehensive school	4963
6	Vocational school	457
7	Other school	2263
8	Special school	277
9	Other school (East 90,91)	1732
-1	No answer	3
-2	Does not apply	79640
-3	Answer improbable	61
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

In order to integrate various information on general school education from all waves (A ...BE) into one variable **ks\_gen**, the different codes of different questionnaire items were modified. In 1992, upper secondary schooling (grammar school) received a new value, thus a recode was necessary to combine information of waves A ...H and I ...K. In 1995, vocational school becomes an item to general education in the wave-specific files **sKIND**. This change for waves A ...H and I ...K is visible in **ks\_gen** since value 6 and 7 for years 1984-1994 are also generated separately in variable **ks\_voc**. For waves G and H East German school attendance was surveyed separately. This information can be found in **ks\_gen** represented by 9 and also in variable **ks\_east**. For the remaining waves L ...BE the answers have the above depicted values.

### **ks\_vocat** – Vocational school

---

1	Vocational school	231
2	Other school	192
-1	No answer	1
-2	Does not apply	173519
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **ks\_vocat** describes vocational school participation in addition to **ks\_gen** in order to depict the different instruments used in waves A ...K and L ...Z. Between wave H and I the values for vocational school training were modified and hence **ks\_vocat** in-

cludes a recode. From 1984 until 1991 vocational school was surveyed using more items to describe different types of vocational schooling. These 4 types are embedded in value 1 of **ks\_vocat**. This information can be found in **ks\_gen** in values 6 and 7. Hence **ks\_vocat** only comprises waves A ...K.

### **ks\_none** – None of these institutions

---

1	None of these institutions (1984-2008)	19016
2	None of these schools (2009)	17379
3	No pre-school attended (2009)	3947
-1	No answer	5628
-2	Does not apply	127970
-3	Answer improbable	3
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **ks\_none** collects missing information for the variables **ks\_pre** and **ks\_gen**. For **ks\_none** value 1 describes whether children do not attend preschool **ks\_pre** or do not attend any general schools depicted in **ks\_gen**. From 1984 until 2008 (waves A ...Y), value 1 summarizes if none of these institutions were attended. In 2009 (wave Z) the survey mode changed: preschool attendance is inquired separately from **ks\_gen** and this new variable includes non-participation. Value 3 of **ks\_none** depicts this change. Also value 2 (none of these schools) indicates that for wave Z the old variable now contains only non-participation of general school education. Values 2 and 3 apply to waves Z ...BE.

### **ks\_east** – General school east germany

---

1	Polytechnic Oberschule	1618
2	Extended Oberschule	31
3	Specialized school	18
4	Special school	27
5	Other comprehensive school	15
6	Vocational school	10
7	Other school	13
-1	No answer	0
-2	Does not apply	172211
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **ks\_east** summarizes additional school type information collected in waves G and H (1990/ 91) in East Germany. The observations are also embedded in **ks\_gen** value 9.

### **ks\_eastp** – Pre-school east germany

---

1	Boarding school (East 90/ 91)	369
---	-------------------------------	-----



-1	No answer	0
-2	Does not apply	173574
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **ks\_eastp** summarizes additional preschool information collected in waves G and H (1990/ 91) in East Germany. The observations are also embedded in **ks\_pre**.

### **ks\_other** – Other type of school

---

1	Special school	15
6	Learning disabilities school	1
7	Remedial school	5
10	School for the handicapped	4
11	School for the blind	1
16	School for disabled	1
17	Observation period	6
20	Speech therapy school	1
21	Speech therapy primary school	1
22	Speech therapy day center	1
24	Business school	2
25	Continuing education	1
26	Vocational school	1
27	Vocational school continuing education	1
28	Professional preparation year	1
29	Domestic science school	1
31	Home economics school	1
32	Special assistance school	1
34	School for the hearing-impaired	1
35	School for learning disorders	3
36	Special school (run by church)	2
37	Special school for learning disabilities	1
-1	No answer	19
-2	Does not apply	173872
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **ks\_other** summarizes additional school type information collected in wave L and M (1995 /96). The observations are also embedded in **ks\_gen** in value 7.

## **6 Variables on day care arrangements (var names starting with KD)**

### **kd\_publ** – Provider: public facility

---

1	Public provider	18061
---	-----------------	-------

2	Public school (2011)	10248
3	Public provider (2011)	3758
-1	No answer	213
-2	Does not apply	141663
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kd\_publ** summarizes information collected in waves D, L, N, S, V, X and BB. In 1986, provider of facility was surveyed for the first time and did not comprise an item for private provider. The information from waves D and L were integrated in the variables format of waves N, S, V and X. This format is used in **kd\_publ** - **kd\_comm**. In 2011 provider of facility was separately inquired for institutions of general school education **ks\_gen** and day care institutions **ks\_pre**. To depict this change two additional values were introduced to all variables **ks\_publ**, **kd\_indep**, **kd\_priv**, **kd\_comp** and **kd\_comm**. The variable **kd\_comm** and **kd\_publ** describe publicly provided facilities and value 2 depicts only public schools and value 3 only public day care institutions.

#### **kd\_priv** – Provider: church facility

---

1	Private provider	214
2	Private non-profit school (2011)	215
3	Private provider (2011)	553
-1	No answer	0
-2	Does not apply	172961
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kd\_priv** summarizes information collected in waves D, L, N, S, V, X and BB. The variable **kd\_priv** describes day care facilities provided by private non-profit organizations. Value 2 depicts only non-profit private schools and value 3 only non-profit private day care institutions.

#### **kd\_indep** – Provider: private non-profit facility

---

1	Independent provider	2592
2	Church school (2011)	641
3	Church instituion (2011)	3259
-1	No answer	213
-2	Does not apply	167238
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kd\_indep** summarizes information collected in waves D, L, N, S, V, X and BB. The variable **kd\_indep** describes day care facilities provided by independent organi-

zations, e.g., churches. Value 2 depicts only independent schools and value 3 only independent day care institutions.

#### **kd\_comp** – Provider: company facility

---

1	Provided by company	28
3	Provided by company (2011)	47
-1	No answer	213
-2	Does not apply	173655
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kd\_comp** summarizes information collected in waves D, L, N, S, V, X and BB. The variable **kd\_comp** describes day care facilities provided by companies. Value 3 depicts day care institutions provided by companies in 2011.

#### **kd\_comm** – Provider: private-commercial facility

---

1	Commercial provider	354
2	Private school (2011)	119
3	Commercial provider (2011)	358
-1	No answer	543
-2	Does not apply	172569
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kd\_comm** summarizes information collected in waves D, L, N, S, V, X and BB. The variable **kd\_comm** describes day care facilities provided by commercial organizations. Value 2 depicts only private schools and value 3 only commercial day care institutions.

#### **kd\_rely** – Adherence to opening time

---

1	Fully rely	4210
2	Partly rely	661
3	Cannot rely	93
-1	No answer	151
-2	Does not apply	168828
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kd\_rely** was only collected in wave S. It comprises information on the adherence to opening times of day care facilities.

**kd\_time** – Time spent at facility

---

1	In the morning	45803
2	In the afternoon	390
3	Full-time	18745
4	No, not full-time (2009)	27399
-1	No answer	1808
-2	Does not apply	79798
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kd\_time** summarizes at what time during the day children attend a day care institution. The institutions are pre-school, creche and nursery (also after school nursery). The variable **kd\_time** is observed from wave L ...BB. Throughout this time period the survey mode changed. In 2009 the question specifically asks whether a child attends the school full time or not (see **ks\_gen**). For waves L ...Y the options are represented by values 1, 2, and 3 and value 4 depicts answers since wave Z.

**kd\_hrs** – Hours p.day spent at facility

---

-1	No answer	57
-2	Does not apply	152306
-3	Answer improbable	10
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kd\_hrs** contains information on how many hours per day a child attends an institution. The institutions are pre-school, creche and nursery (also after school nursery). The variable **kd\_hrs** has been introduced in the SOEP in 2009.

**kd\_lunch** – Lunch at facility

---

1	Yes, lunch provided	12644
2	No lunch provided	12209
-1	No answer	369
-2	Does not apply	148721
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kd\_lunch** contains information collected in waves N, S, V, X and BB. It summarizes whether a child that attends a day care facility receives a lunch at the facility.

**7 Variables on cost of school (var names starting with KS)**

**ks\_lunch** – Lunch at school

---

1	Yes	4942
2	No	3076
-1	No answer	34
-2	Does not apply	165891
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

**ks\_amtp** – Costs for school per month

---

-1	No answer	0
-2	Does not apply	172009
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

**ks\_cost** – School: no costs

---

1	Incurred no costs	6060
-1	No answer	58
-2	Does not apply	167825
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

**8 Variables on costs of pre-school (var names starting with KS)****kk\_amtp** – Costs for pre-school per month

---

-1	No answer	0
-2	Does not apply	170279
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

**kk\_cost** – Pre-school: no costs

---

1	Incurred no costs	880
-1	No answer	37
-2	Does not apply	173026
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0

-6 Version of questionnaire with modified filtering 0

## 9 Variables on cost of day care arrangements (var names starting with KD)

### kd\_fee – Payment of fee

---

1	Free of charge	12923
2	Fixed sum in Euro	5246
-1	No answer	308
-2	Does not apply	155466
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kd\_fee** summarizes observations of waves D, M, S, V, and X. It comprises information on whether parents pay a day care fee or not.

### kd\_amtp – Amount paid

---

-1	No answer	170
-2	Does not apply	168653
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

If parents pay a fixed sum in EURO **kd\_amtp** comprises the actual amount paid. In waves D and M the amount was indicated in DM (Deutsche Mark). These amounts have been converted to Euro (1 Euro = 1.95593 DM).

## 10 Variables describing carer outside the household (var names starting with KC)

### kc\_relat – Carer: relatives

---

1	Carer: relatives	22188
-1	No answer	1
-2	Does not apply	151754
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kc\_relat** summarizes information collected in wave D, N ...S, U ...BE. In 1986, help with care was surveyed using one variable and it does not contain information on child minder. From wave N onwards carers became a separate group in the household

questionnaire yet the item child minder was only integrated into carers in 2005. Variable **kc\_relat** comprises information available from waves D, N ...S and U ...BE observing that a child is in informal care provided by relatives.

### **kc\_frnd** – Carer: friends

---

1	Carer: friends	3850
-1	No answer	1
-2	Does not apply	170092
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kc\_frnd** summarizes information collected in wave D, N ...S, U ...BE. The variable **kc\_frnd** comprises information available from waves D, N ...S and U ...BE observing that a child is in informal care provided by friends, neighbors, etc..

### **kc\_paid** – Carer: paid

---

1	Carer: paid care	4056
-1	No answer	1
-2	Does not apply	169886
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kc\_paid** summarizes information collected in wave D, N ...S, U ...BE. The variable **kc\_paid** comprises information available from waves D, N ...S and U ...BE observing that a child is in informal care provided by paid carers, i.e. nannies and babysitters.

### **kc\_mindr** – Carer: child minder

---

1	Carer: child minder	1846
-1	No answer	0
-2	Does not apply	172097
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kc\_mindr** summarizes information collected in wave D, and L ...BE. In 1986, help with care was surveyed using one variable and it does not contain information on child minder. This was introduced in 1995 as alternative item to pre-school attendance (L ...P). From wave N onwards carers became a separate group in the household questionnaire yet the item child minder was only integrated into carers in 2005. Before wave Q ...U child minder remained grouped in institutions attended. The variable **kc\_mindr** comprises information available from waves D and L ...BE observing that a child is in care

provided by a child minder.

### kc\_none – Carer: none

---

1	Carer: none	85599
-1	No answer	3434
-2	Does not apply	84910
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

The variable **kc\_none** collects the missing information for variables **kc\_relat** - **kc\_mindr**. It comprises all information available observing that a child is not using informal care.

## 11 Activities below the age of 6 (var names starting with KA06)

The variables **ka06\_spo** - **ka06\_oth** summarize activities for children below the age of 6. This information has been surveyed for the first time in 2006 and is supposed to be collected every other year. Due to the integration of the FID-samples this information is also available for waves BB and BD. Hence so far the following variables are available for waves W, Y and BE.

### ka06\_spo – Gymnastics under 6

---

1	Yes, activity applies	8341
-1	No answer	0
-2	Does not apply	155098
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

### ka06\_mus – Music education under 6

---

1	Yes, activity applies	3271
-1	No answer	0
-2	Does not apply	160168
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

### ka06\_art – Art class under 6

---

1	Yes, activity applies	769
-1	No answer	0
-2	Does not apply	162670



-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

### ka06\_oth – Other activities under 6

---

1	Yes, activity applies	3258
-1	No answer	17
-2	Does not apply	160164
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

### ka06\_non – No activities under 6

---

1	No activities under 6	11632
-1	No answer	376
-2	Does not apply	157671
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	4264
-6	Version of questionnaire with modified filtering	0

The variable **ka06\_non** indicates that a child does not participate in any activities for children below the age of 6.

## 12 Activities for children aged 6 to 16 (var names starting with KA16)

### ka16\_ssp – School sports club above 6

---

1	Yes, activity applies	7697
-1	No answer	0
-2	Does not apply	155742
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

### ka16\_smu – School music club above 6

---

1	Yes, activity applies	5091
-1	No answer	0
-2	Does not apply	158348
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

**ka16\_sar** – School art club above 6

---

1	Yes, activity applies	1724
-1	No answer	0
-2	Does not apply	161715
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

**ka16\_sth** – School theater club above 6

---

1	Yes, activity applies	1591
-1	No answer	0
-2	Does not apply	161848
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

**ka16\_sot** – Other school activities above 6

---

1	Yes, activity applies	4988
-1	No answer	0
-2	Does not apply	158451
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

**ka16\_spo** – Sports club above 6

---

1	Yes, activity applies	19730
-1	No answer	0
-2	Does not apply	143709
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

**ka16\_mus** – Music education above 6

---

1	Yes, activity applies	8290
-1	No answer	0
-2	Does not apply	155149
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

**ka16\_art** – Art class above 6

---

1	Yes, activity applies	628
-1	No answer	0
-2	Does not apply	162811
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

**ka16\_org** – Volunteer fire dept, red cross etc. Above 6

---

1	Yes, activity applies	1398
-1	No answer	0
-2	Does not apply	162041
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

**ka16\_yth** – Youth group above 6

---

1	Yes, activity applies	2883
-1	No answer	0
-2	Does not apply	160556
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

**ka16\_ctr** – Youth center above 6

---

1	Yes, activity applies	933
-1	No answer	0
-2	Does not apply	162506
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

**ka16\_non** – No activities above 6

---

1	No activities above 6	7115
-1	No answer	312
-2	Does not apply	156012
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	10504
-6	Version of questionnaire with modified filtering	0

The variable **ka16\_non** indicates that a child does not participate in any activities for children above the age of 6.

### 13 Variables on costs of school, carer, activities etc (var names starting with K)

#### **k\_amp** – Costs per month

---

-1	No answer	0
-2	Does not apply	152228
-3	Answer improbable	1
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	16437
-6	Version of questionnaire with modified filtering	0

In wave BA (2010) average monthly costs were surveyed: ‘How much do the above mentioned school, care and activities cost you?’ Parents indicate average monthly cost of their children regarding school, care, and (leisure) activities, e.g., early music education, sports, etc.. In wave BA this variable was surveyed for the first time. Other than variable **kd\_amp**, **k\_amp** identifies overall average monthly costs and not just day care fees.

#### **k\_cost** – No costs

---

1	No, costs	3992
-1	No answer	572
-2	Does not apply	161453
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	7926
-6	Version of questionnaire with modified filtering	0

Corresponding to parents’ estimation of average monthly costs **k\_amp** “no costs” indicates if parents do not have any costs.

### 14 Variables on costs of external care (var names starting with KC)

#### **kc\_amp** – External care: costs per month

---

-1	No answer	0
-2	Does not apply	166919
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

In wave BB (2011) another aspect of monthly costs was surveyed: ‘How much do the above mentioned care arrangements outside the household cost?’ Parents indicate average monthly cost of their children regarding additional care outside the household, e.g.,

relatives, friends, neighbors, etc..

**kc\_cost** – External care: no costs

---

1	No costs	8399
-1	No answer	176
-2	Does not apply	165368
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0

Corresponding to parents' estimation of average monthly costs "no costs" indicates if parents do not have any costs.