Documentation for the Standardization of the Bulgarian Harmonized Histories Data File for birth, partnership histories, leaving home questions and background variables

HARMONIZED HISTORIES BULGARIA (12858 respondents)

Karolin Kubisch
Max Planck Institute for Demographic Research Rostock

2014
Updated 27.10.2015

The following documentation gives a description of all input variables and the consequent preparation of the output variables according to the manual for the preparation of comparative fertility and union histories.
All problem cases as well as the treatment of these cases are described in detail. At the end of each module a summary of the main findings is displayed (in red).

Missing values are coded:
.a unknown
.b does not apply
.c unavailable in survey

Source: GGS first wave, GGS_Wave1_Bulgaria_V.4.0.dta

Interview dates Bulgaria GGS First wave: October to December 2004

June 2014: Corrections in the variables to leaving home histories of children (KID_L, KID_LY, KID_LM)

October 2015: Please note that the partnership histories were modified in October 2015. More precisely, we changed the sorting of the variable UNION_$ (Union order). Prior to that date, we had sorted the unions by start year of the union. This involved that unions which start dates were missing were always listed as last unions. In the modified version, we sorted the partnerships no longer by relying on the start year of the union, but by relying on the order of the union as they appear in the original dataset. For Bulgaria it affects ca. 10 cases.

In connection with this modification, some smaller consistency changes were made to the data. In particular, we recoded the following constellations:

- Events (Union, Marriage, Separation, Divorce) before age 12 of respondent
- Event before age 12 of partner
- Negative difference between partnership date and marriage date
- Negative difference between separation date and union or marrige date and negative difference between divorce date and union or marriage date

- Sucessive partnerships mar-mar[_n-1]<=0 or par-par[_n-1]<=0
- Differences between separation date and next partnership date sep>par[_n+1]

All modifications made October 2015 are described in the updated documentation.

---

1. Part Basic Information

RESPID: ID number to be assigned at merging — LEAVE BLANK

ARID: ID number from raw data (original ID number) — used: arid
12858 respondents

COUNTRY: Country and survey — used: acountry
acountry: code: 1: Bulgaria
COUNTRY: code: 1001: Bulgaria GGS Wave 1
no missing cases

MONTH_S: Month of survey — used: amonth
amonth: codes: 10-12
43 former missing cases(in UN version recoded as june) were recoded to .a

IMONTH_S: Month of survey, including imputed dates — used: amonth
For missing values imputation:
randomly variable between 10 and 12

YEAR_S: Year of survey — used: ayear
ayear: 2004
YEAR_S: 2004
43 missing cases
➔ Imputation: 2004

SEX: Sex of the respondent — used: ahg4_1
No missing cases
Sex structure of the Bulgarian respondents:
Male: 5851 and Female: 7007

BORN_Y: Year of birth of respondent — used: ahg6y_1
ahg6y_1: 1919-1987: 3 missing cases

BORN_M: Month of birth of respondent — used: ahg6m_1
63 missing cases + additionally 5 seasonal codes

IBORN_M: Month of birth of respondent — used: BORN_M
including imputed months
randomly variable between 1-12

---

2. Part LEAVING HOME

LEAVE_1: Indicator of whether “left home”
Used: GRID=1 go to a5117a
GRID=0 go to a5116m/y
a5117a=1 go to a5117bm/y

Definition:
* Respondent did not leave home (code 0) if: a parent lives in the household (GRID=1) and respondent never lived separately from parents (a5117a=2)
* Respondent left home (code 1) if: there is no parent in household (GRID=0) or there is a parent in household (GRID=1) and respondent ever left home (a5117a=1)

LEAVE_1: 0: 2053 / 1: 10727
78 missing cases

LEAVE_Y1: Year of first time leaving home
used: a5116y and a5117by

Filter: LEAVE_Y1/LEAVE_M1: Transformation to .b (Does not apply)if LEAVE_1==0 (2053)
Missing cases: .b 2053 .a 326

LEAVE_M1: Month of first time leaving home
used: a5116m and a5117bm
LEAVE_M1: codes: 1-12 and additionally seasonal codes
Missing cases: .b 2053 .a 1235

ILEAVE_M1: Month of first time leaving home and imputed months: used: LEAVE_M1

Harmonized: random variables according to manual
Filter: .b 2053

---

3. Part UNIONS AND DISSOLUTION ($=order of union)

UNINUM: Total number of unions used: UNION_1 to _4

Syntax:
```
forvalues x=1/4 {
replace UNINUM=UNINUM+1 if UNION_`x'>0
            }
```
UNINUM:
0: 2969
1: 9309
2: 552
3: 26
4: 2

UNION_$: UNION order

For the chapters union /marriage and divorce/ and a part of partners characteristics an reshaping program was used, which includes partnership histories and questions to the current partner

Definition (UNION_1 to UNION_x)
➔a union exists if there is an answer in at least one of the questions about the current partner (a301m – a309) or in partnership histories (a334m – a350)

UNION_1: 9889 .d 152
UNION_2: 580 .d 8
UNION_3: 28
UNION_4: 2

UNION_Y$: Year of start union
used: a301y and a334y

Filter: UNION_Yx=.b if UNION_x==0

UNION_Y1 missing values: 67
UNION_Y2 missing values: 22
UNION_Y3 missing values: 2
UNION_Y4 missing values: 1

Problems and transformations connected with partnerships

TRANSFORMATIONS:

```
replace a301y=1984 if ARID==43160
replace a301y=.a if ARID==133379 | ARID==204196 | ARID==765807 |
ARID==927930 | ARID==1896506 | ARID==2196403 | ARID==3167736 |
ARID==5671562 | ARID==6396259 | ARID==8275369 | ARID==8497459
replace a302by=.a if ARID==133379 | ARID==204196 | ARID==927930 |
ARID==8497459
replace a301y=1967 if ARID==1323359
replace a301y=1981 if ARID==2120386
replace a301y=1991 if ARID==3567459
replace a301y=1992 if ARID==8254012
replace a302by=1982 if ARID==3383345
replace a302by=.a if ARID==6097098
replace a302by=2001 if ARID==9992176
replace ahg6y_2=.a if ARID==105308 | ARID==2078382 | ARID==71980198 |
ARID==9447227
replace ahg6y_2=.a if ARID==151711 | ARID==2266293 | ARID==2653340 |
ARID==3642155 | ARID==5878366 | ARID==7198019
replace a301y=.a if ARID==949851
replace a301y=.a if ARID==72994 | ARID==7065826 | ARID==7812382 |
ARID==7855701
replace a301y=.a if ARID==229041 | ARID==1782600 | ARID==3483261 |
ARID==3762729 | ARID==6045385 | ARID==6075518 | ARID==6699065 |
ARID==7156028 | ARID==1005256 | ARID==7235779 | ARID==7249404 |
ARID==7376278 | ARID==8572352 | ARID==9316126
replace a301m=9 if ARID==1675040
replace a301m=12 if ARID==6616961
replace a302by=.a if ARID==6699065 | ARID==1005256
```

```
replace a334y_1=.a if ARID==339509 | ARID==5074152 | ARID==5288120 |
ARID==5761088 | ARID==6577811 | ARID==7242170 | ARID==7646055 |
ARID==9144697 | ARID==9833558
replace a335y_1=.a if ARID==5074152 | ARID==7242170 | ARID==7646055
replace a334y_1=1949 if ARID==2546024
replace a334y_1=1989 if ARID==6105132
replace a334y_1=1996 if ARID==7334288
replace a335y_1=1955 if ARID==529422
replace a335y_1=.a if ARID==2846925 | ARID==4486698
replace a349y_1=.a if ARID==8830895
replace a349y_2=.a if ARID==1591050
replace a336y_1=.a if ARID==2840146 | ARID==4866170 | ARID==5860841 |
ARID==8560591 | ARID==9099155 | ARID==622418 | ARID==3803129 |
ARID==6234696 | ARID==8063432 | ARID==8128041 | ARID==8529785
replace a336y_1=.a if ARID==1981 | ARID==770193 | ARID==2103215 |
ARID==2556383 | ARID==3453818 | ARID==9327501 | ARID==9557011 |
ARID==9953714 | ARID==8770529
replace a334y_1=.a if ARID==408783
replace a336y_2=.a if ARID==9696686
replace a335y_1=.a if ARID==350769 | ARID==3571194 | ARID==6452890 |
ARID==6699317
replace a344y_1=.a if ARID==5604884 | ARID==5970592 | ARID==8268729 |
ARID==9046294 | ARID==9232347
replace a349y_1=.a if ARID==8268729 | ARID==2261586
replace a344y_1=1997 if ARID==8529785
replace a334y_2=. if ARID==229572 | ARID==5908333 | ARID==6810848 |
ARID==3956465 | ARID==7734332 | ARID==3593409 | ARID==7044090 |
ARID==5761088
replace a334m_2=. if ARID==229572 | ARID==5908333 | ARID==6810848 |
ARID==3956465 | ARID==7734332
replace a335a_2=. if ARID==229572 | ARID==5908333 | ARID==6810848 |
ARID==3956465 | ARID==7734332
replace a336m_2=. if ARID==229572 | ARID==5908333 | ARID==6810848 |
ARID==3956465 | ARID==7734332
replace a336y_2=. if ARID==229572 | ARID==5908333 | ARID==6810848 |
ARID==3956465 | ARID==7734332
replace a337_2=. if ARID==229572 | ARID==5908333 | ARID==6810848 |
ARID==3956465 | ARID==7734332
replace a338_2=. if ARID==229572 | ARID==5908333 | ARID==6810848 |
ARID==3956465 | ARID==7734332
replace a343_2=. if ARID==229572 | ARID==5908333 | ARID==6810848 |
ARID==3956465 | ARID==7734332
replace a344m_2=. if ARID==229572 | ARID==5908333 | ARID==6810848 |
ARID==3956465 | ARID==7734332
replace a344y_2=. if ARID==229572 | ARID==5908333 | ARID==6810848 |
ARID==3956465 | ARID==7734332 | ARID==3593409 | ARID==5761088
replace a345_2=. if ARID==3956465
replace a335y_2=. if ARID==780443 | ARID==5761088 | ARID==3593409
replace a335m_2=. if ARID==5908333 | ARID==6810848 | ARID==7734332
replace a335y_2=. if ARID==5908333 | ARID==6810848 | ARID==7734332
replace a335y_1=1967 if ARID==7368241
replace a335y_2=2000 if ARID==8474141
replace a334m_3=. if ARID==9115534
replace a334y_3=. if ARID==9115534
replace a335a_3=. if ARID==9115534
replace a336m_3=. if ARID==9115534
replace a336y_3=. if ARID==9115534
```

```
replace a337_3=. if ARID==9115534
replace a338_3=. if ARID==9115534
replace a343_3=. if ARID==9115534
replace a344m_3=. if ARID==9115534
replace a344y_3=. if ARID==9115534
replace a344y_1=.a if ARID==5067945 | ARID==5623201 | ARID==6810848
replace a334m_2=10 if ARID==2289245
replace a334m_2=7 if ARID==4506662
replace a344y_1=1983 if ARID==7109267
replace a334y_1=1989 if ARID==8365413
```

UNION_M$: Month of start UNION used: a301m and a334m

Filter: UNION_Mx=.b if UNION_x==0

UNION_M1 missing values: 588 + additional seasonal codes
UNION_M2 missing values: 40 + additional seasonal codes
UNION_M3 missing values: 3 + additional seasonal code

IUNION_M$: Month of start UNION and imputed months according to manual page 4 (random) used: UNION_M$

Filter: _IUNION_Mx=.b if UNION_x==0

Summary "UNION":

Some problems with start dates of the union were found and some transformations had to be performed which are described in the chapter above.

SEP_$: Dissolution of UNION used: a343 (only histories)

Filter: SEP_x=.b if UNION_x==0

➔ in case of current partner: no separation
SEP_1 missing cases: 14
SEP_2 missing cases: 5
SEP_3 missing cases: 1

Order of Union	Number of unions	number of separations	death of partner
1	9889	1045	760
2	580	90	25
3	28	6	
4	2	1	

SEP_Y$: Year of end of UNION used: a344y (only for histories)

Filter: SEP_Yx=.b if UNION_x==0
SEP_Yx=.b if SEP_x==0

SEP_Y1 missing values: 97
SEP_Y2 missing values: 11
SEP_Y3 missing values: 1

SEP_M$: Month of end of UNION used: a344m (histories only)

Filter: SEP_Mx=.b if UNION_x==0
SEP_Mx=.b if SEP_x==0

SEP_M1 missing values: 297 + additional seasonal codes
SEP_M2 missing values: 23 + additional seasonal codes
SEP_M3 missing values: 2

ISEP_M$: Month of end of UNION and imputed months according to manual page 4 (random) used: SEP_M$

Filter: ISEP_Mx=.b if UNION_x==0
ISEP_Mx=.b if SEP_x==0

Summary "Separation":
Some problems with dates of the separation were found and some transformations had to be performed which are described in the chapter above.

---

4. Part MARRIAGE AND DIVORCE ($=order of union)

MARR_$: Indicator of whether marriage took place and type of marriage used: a302a and a335a

Filter: MARR_x=.b if UNION_x==0
MARR_1 missing values: 81
MARR_2 missing values: 8
MARR_3 missing values: 1

Order of Union	Number of unions	number of marriages
1	9889	8912
2	580	292
3	28	11
4	2	

MARR_Y$: Year of marriage used: a302by and a335y

Filter: MARR_Yx=.b if UNION_x==0
MARR_Yx=.b if MARR_x==0

MARR_Y1 missing values: 166
MARR_Y2 missing values: 19
MARR_Y3 missing values: 1

MARR_M$: Month of marriage used: a302bm and a335m

Filter: MARR_Mx=.b if UNION_x==0

MARR_Mx=.b if MARR_x==0

MARR_M1 missing values: 517 + additional seasonal codes
MARR_M2 missing values: 32 + additional seasonal codes
MARR_M3 missing values: 1

IMARR_M$: Month of marriage and imputed months according to manual page 4 (random) — used: MARR_M$

Filter: IMARR_Mx=.b if UNION_x==0
IMARR_Mx=.b if MARR_x==0

Summary "Marriage":

Some problems with dates of the marriage were found and some transformations had to be performed which are described in the chapter above. Some problematical cases remain.

DIV_$: Indicator of whether divorce occurred — used: a349a, a343 (only histories)

Filter: DIV_x=.b if UNION_x==0
DIV_x=.b if MARR_x==0
DIV_x=.d if a343_x==2

DIV_1 missing values: 44
DIV_2 missing values: 9

Order of Union	Number of unions	number of marriages	number of divorces
1	9889	8912	762
2	580	292	35
3	28	11	2
4	2		

DIV_Y$: Year of divorce — used: a349y

Filter: DIV_Yx=.b if UNION_x==0
DIV_Yx=.b if MARR_x==0
DIV_Yx=.b if DIV_X==0 or .d

DIV_Y1 missing values: 77
DIV_Y2 missing values: 12

DIV_M$: Month of divorce — used: a349m

Filter: DIV_Mx=.b if UNION_x==0
DIV_Mx=.b if MARR_x==0
DIV_Mx=.b if DIV_x==0 or .d

DIV_M1 missing values: 184 + additional seasonal codes
DIV_M2 missing values: 18

IDIV_M$: Month of divorce and imputed months according to manual page 4 (random) used: DIV_M$

Filter: IDIV_Mx=.b if UNION_x==0
IDIV_Mx=.b if MARR_x==0
IDIV_Mx=.b if DIV_x==0 or .d

Summary "Divorce":
Some problems with dates of the divorce were found and some transformations had to be performed which are described in the chapter above.

---

5. Part PARTNER`S CHARACTERISTICS ($=order of union)

SEXP_$: Partner`s sex used: ahg4_2, ahg4_1, a352a

For current partnership: ahg4_2
For histories: a352a (homosexual partnership): 1 case

Filter: SEXP_x=.b if UNION_x==0

Partner	Number of unions	Number male	Number female
1	9889	5603	4286
2	580	328	252
3	28	12	16
4	2		2

YEARBIRP_$: Year of birth of partner Used: ahg6y_2 and a336y

Filter: YEARBIRP_x=.b if UNION_x==0

YEARBIRP_1 missing cases: 211
YEARBIRP_2 missing cases: 14
YEARBIRP_3 missing cases: 1

MONBIRP_$: Month of birth of partner used: ahg6m_2 and a336m

Filter: MONBIRP_x=.b if UNION_x==0

MONBIRP_1 missing cases: 422 + additional seasonal codes
MONBIRP_2 missing cases: 32
MONBIRP_3 missing cases: 42

IMONBIRP_$: Month of birth of partner and imputed months according to manual page 4 (random) used: MONBIRP_$

Filter: IMONBIRP_x=.b if UNION_x==0

NUMCHP_$: Number of children of partner at start of union$

for current partner:
a)children of partner (household members): relation of household member to respondent (ahg3_): code 4: stepchild: my current partner's children not adopted by me (128 children)➔ ahg3_3 to ahg3_8
b)non-resident stepchildren: a226==1 (yes: 253) and a229
c)for partnership histories: a338_1 to a338_8
also: year of start of union(a301y) and year of birth of stepchild (ahg6y_x and a230_x)

Problem: The question: When did you start living together, how many children did your partner have? (a338)- exists only for partnership histories
-for current partnership it had to be created with the help of the number of stepchildren, year of start of union and year of birth of stepchild

Definition:
the number of children of current partner includes:
* all stepchildren of respondent living at the moment of the interview in household grid and were born before the start of the union
* all nonresident stepchildren at the time of interview – partner's children born before partnership (year start union - birth year>0)
* the number of partner`s children at start of a union in partnership history (a338_1 to a338_8)

Filter: NUMCHP_$=.b if UNION_X==0

NUMCHP_1: missing values: 338
NUMCHP_2: missing values: 20
NUMCHP_3: missing values: 1

NUMCLIV_$: Number of children of partner lived with respondent

Problem: The question: How many of them lived with respondent (a341)- exists only for partnership histories.
➔for current partnership it had to be created

Definition:
the number of children of current partner ever lived with respondent includes:
* all stepchildren of respondent living at the moment of interview in household grid
* all nonresident stepchildren at the time of interview (partner`s children born before partnership), who ever lived in respondent's household for more than 3 months (a231_1 to a231_8)
* the number of partner`s children, who lived with respondent in a union in partnership history (a341_1 to a341_8)

NUMCLIV_1: missing values: 343
NUMCLIV_2: missing values: 23
NUMCLIV_3: missing values: 1

Union	Number of unions	NUMCHP	NUMCLIV
1	9884	1:89 2:58 3:6 4:3 9:1	1:78 2:49 3:6 4:4 6:1
2	583	1:121 2:75 3:12 4:5 5:2 6:2	1:106 2:69 3:11 4:3 5:4 6:2
3	30	1:6 2:6 4:1 5:1	1:6 2:7 5:1
4	3	1:1	

Summary :
The variables NUMCHP and NUMCLIV had to be created for the current partnership.

6. Part Birth histories (biological kids)

For the chapter “Birth histories” a reshaping program was used, which includes questions to the biological children in the household and to the nonresident biological children

To create the number of biological children (KID_1 to KID_x)the following definition was applied:
➔a biological child exists in a household if there is code 2 or 3 (biological child by current or previous partner) in the relationship to respondent (ahg3_)
➔a nonresident biological child exists if a213_==1

KID_$: Indicator of child order

used: ahg1_ and generated variable obnr (at least 1 answer in questions a212 to a224)
no missing cases

Child order	number of children	.d
1	9472	51
2	6253	100
3	1067	104
4	303	23
5	121	9
6	55	3
7	22	3
8	14	2
9	6	
10	1	
11	1	
12	1	

KID_Y$: Year of birth of child used: ahg6y_ and a216y

Filter: KID_Yx=.b if KID_x==0
KID_Y1 missing values: 133
KID_Y2 missing values: 183
KID_Y3 missing values: 79
KID_Y4 missing values: 25
KID_Y5 missing values: 17
KID_Y6 missing values: 8
KID_Y7 missing values: 4
KID_Y8 missing values: 3
KID_Y9 missing values: 2

Problems with year of birth of child:

TRANSFORMATIONS

```
replace ahg6y_3=.a if ARID==204196
replace ahg6y_3=.a if ARID==1812921 | ARID==1975226 | ARID==2318040 |
ARID==4498866 | ARID==4779026 | ARID==7504165 | ARID==9226364 |
ARID==3167736
replace ahg6y_4=.a if ARID==4779026

replace a216y_1=1974 if ARID==1519243
replace a216y_1=.a if ARID==2116481 | ARID==665555 | ARID==3744747 |
ARID==4530892 | ARID==5543297 | ARID==6257261 | ARID==6418155 |
ARID==6447930 | ARID==6696211 | ARID==7519018 | ARID==8083726 |
ARID==9466983 | ARID==9720980
replace a220y_1=.a if ARID==1366723 | ARID==3238744  | ARID==3655455 |
ARID==5990778 | ARID==8338483 | ARID==8357578 | ARID==690565 |
ARID==2452283 | ARID==2625605 | ARID==2950568 | ARID==3769378 |
ARID==7735956 | ARID==7885635 | ARID==7969573 | ARID==8995644 |
ARID==9797094
replace a220y_2=.a if ARID==6721614 | ARID==1320788 | ARID==1551977 |
ARID==2304191 | ARID==2919201 | ARID==2950568 | ARID==3749142 |
ARID==3769378 | ARID==3898382 | ARID==4013390 | ARID==5389329 |
ARID==5973729 | ARID==8995644 | ARID==9720980
replace a220y_3=.a if ARID==1981 | ARID==1449725
replace a216y_2=.a if ARID==665555
```

Interval between two births < 7months or >20 years:
KID1-KID_2

ARID	SEX	BORN_Y	KID_Y1	KID_M1	KID_Y2	KID_M2
510421	Male	1932	1962	September	1963	February
812394	Female	1954	1975	June	1975	July
829448	Female	1966	1984	September	2004	October
2283360	Male	1956	1971	July	1971	November
2304191	Male	1937	1964	December	2003	June
2550563	Female	1960	1979	December	1980	August
2975854	Female	1941	1963	June	1963	December
3023674	Male	1927	1951	February	1951	April
3383101	Male	1940	1965	July	1966	January
3744747	Female	1935	1951	February	1972	July
3902484	Male	1942	1964	September	1965	February
4013390	Male	1928	1948	January	2002	March
4433072	Female	1956	1981	May	1982	January
4828598	Female	1946	1965	January	1965	September

4972096	Female	1964	1985	March	1985	June
4977337	Male	1925	1954	November	1955	May
5222684	Male	1950	1966	October	1967	May
5866672	Female	1948	1976	July	1977	January
6036119	Male	1936	1964	April	1964	October
6065592	Male	1964	1989	July	1989	December
6676369	Female	1965	1981	October	1982	May
6754911	Male	1970	1996	September	1997	April
6763340	Female	1949	1977	September	1977	December
6793313	Female	1937	1959	August	1960	January
6853873	Female	1973	1990	November	1991	March
7136973	Male	1947	1976	April	1976	May
7223690	Male	1964	1999	September	2000	January
7468053	Male	1941	1969	February	2003	September
7659052	Male	1936	1958	April	1958	July
7743977	Female	1951	1974	January	1974	June
8357578	Male	1947	1970	April	1970	August
8443891	Male	1933	1953	June	1977	October
8565072	Female	1982	2001	October	2002	May
9270982	Male	1934	1953	September	1975	February
9454667	Male	1971	1988	October	1989	March

KID2-KID_3

ARID	KID_Y2	KID_M2	KID_Y3	KID_M3	SEX	BORN_Y
867076	1960	April	1960	July	Female	1936
1878049	1980	September	1981	May	Male	1952
2247357	1971	October	1972	April	Male	1941
3598313	1959	September	1982	February	Male	1934
3716904	1975	April	1997	August	Female	1952
3841910	1972	December	1973	March	Male	1950
5973729	1989	May	1989	July	Male	1957
7362654	1969	July	1992	January	Female	1948

KID3-KID_4

ARID	KID_Y3	KID_M3	KID_Y4	KID_M4	SEX	BORN_Y
6624558	1961	June	1961	October	Female	1936
7004646	1982	September	2002	December	Female	1960

➔NO CHANGES, Only for your information

KID_M$: Month of birth of child used: ahg6m and a216m

Filter: KID_Mx=.b if KID_x==0

KID_M1 missing values: 310+seasonal codes
KID_M2 missing values: 336+seasonal codes
KID_M3 missing values: 141+seasonal codes
KID_M4 missing values: 64+seasonal codes
KID_M5 missing values: 36+seasonal codes
KID_M6 missing values: 18+seasonal codes
KID_M7 missing values: 7
KID_M8 missing values: 8
KID_M9 missing values: 4

IKID_M$: Month of birth of child and imputed months used: KID_M$
according to manual page 4 (random)

Filter: IKID_M_x=.b if KID_x==0

KID_S$: Sex of child used: ahg4 and a212

Filter: KID_Sx=.b if KID_x==0

KID_S1 missing cases: 6
KID_S2 missing cases: 5
KID_S3 missing cases: 1

Child order	number of children	male	female
1	9472	4958	4508
2	6253	3155	3093
3	1067	573	493
4	303	143	160
5	121	55	66
6	55	24	31
7	22	12	10
8	14	9	5
9	6	5	1
10	1	1	
11	1		1
12	1		1

KID_D$: Death of child used: a211b
Filter: KID_Dx=.b if KID_x==0

No missing cases

Child order	number of children	death
1	9472	102
2	6253	116
3	1067	44
4	303	12
5	121	8
6	55	2
7	22	4
8	14	1
9	6	2
10	1	
11	1	
12	1	

KID_DY$: Year of death of child used: a217y

Filter: KID_DYx=.b if KID_x==0
KID_DYx=.b if KID_Dx==0

KID_DY1 missing values: 13
KID_DY2 missing values: 16
KID_DY3 missing values: 8
KID_DY4 missing values: 6
KID_DY5 missing value: 1
KID_DY6 missing value: 1

KID_DM$: Month of death of child used: a217m

Filter: KID_DMx=.b if KID_x==0
KID_DMx=.b if KID_Dx==0

KID_DM1 missing values: 28+seasonal codes
KID_DM2 missing values: 28+seasonal codes
KID_DM3 missing values: 12+seasonal codes
KID_DM4 missing values: 8
KID_DM5 missing values: 2
KID_DM6 missing value: 1

IKID_DM$: Month of death of child and imputed months used: KID_DM

according to manual page 4 (random)

Filter: IKID_DMx=.b if KID_x==0
IKID_DMx=.b if KID_Dx==0

KID_L$: Child left home used: a220y/a220m

Child's parental home leave variable (KID_L) was not constructed perfectly as it was created in wide format instead of long. Namely the error occurred assuming that child's order would perfectly match of those living outside the household. More specifically, if child from outside household changes its order (because of preceding foster/adopted or a step child) and in household grid is reported biological child of the same order, then this particular child will be coded as "0" (did not leave home). Furthermore some children living in the household were coded as left home.

Initially both KID_LY (year of child's home leave) and KID_M (month of child's home leave) variables were constructed correctly, however due to reason that KID_L variable serves as filter for both variables then these variables eventually were changed to either ".b" (does not apply) or ".a " (unknown).

Since june 2014 KID_L is constructed in a long format. In addition children which died were excluded from KID_L=1 and are now coded with special missing code .d and KID_LY and KID_LM for dead children is coded as .b.

Definition: Child left home if a220m_x or a220y_x!=.

Filter: KID_Lx=.b if KID_x==0

Child order	number of children	Left home
1	9472	3369
2	6253	2319
3	1067	424
4	303	116
5	121	48
6	55	20
7	22	10
8	14	8
9	6	
10	1	
11	1	

12	1	

KID_LY$: Year child left home used: a220y

Filter: KID_LYx=.b if KID_x==0
KID_LYx=.b if KID_Lx==0

KID_LY1 missing cases: 243
KID_LY2 missing cases: 193
KID_LY3 missing cases: 78
KID_LY4 missing cases: 32
KID_LY5 missing cases: 16
KID_LY6 missing cases: 8
KID_LY7 missing cases: 4
KID_LY8 missing cases: 3

KID_LM$: Month child left home used: a220m

Filter: KID_LMx=.b if KID_x==0
KID_LMx=.b if KID_Lx==0

KID_LM1 missing cases: 614 + additional seasonal codes
KID_LM2 missing cases: 491 + additional seasonal codes
KID_LM3 missing cases: 151 + additional seasonal codes
KID_LM4 missing cases: 49
KID_LM5 missing cases: 25
KID_LM6 missing cases: 14
KID_LM7 missing cases: 6
KID_LM8 missing cases: 5

IKID_LM$: Month of death of child and imputed months used: KID_LM

according to manual page 4 (random variable)

Filter: IKID_LMx=.b if KID_x==0
IKID_LMx=.b if KID_Lx==0

---

7. Part Education

INSCHOOL: Currently studying at the time of interview used: a151

Currently studying: 1040 respondents
Missing cases: 198

EDU_COU: Highest level of education, country specific used: 148

Missing cases: 9

Definition:
The country specific codes include:
* a 3-digit country prefix(100)
* a 1-digit survey code (Bulgarian GGS=1) and

* a 2-digit country specific code for level of education

ISCED_7: Highest level of education
Achieved according to ISCED 1997 used: EDU_COU

Definition:
replace ISCED_7=1 if EDU_COU==100100 | EDU_COU==100101
replace ISCED_7=2 if EDU_COU==100102
replace ISCED_7=3 if EDU_COU==100103
replace ISCED_7=5 if EDU_COU==100105
replace ISCED_7=6 if EDU_COU==100106
replace ISCED_7=.a if ISCED_7==.

Missing cases: 160

Harmonized:

ISCED	Number
0+1	891
2	2693
3	6562
4	
5	2507
6	45

EDU_3: Highest level of education ISCED used: ISCED_7
Collapsed into 3 categories

Definition: High: ISCED_7=code 5 or code 6
Medium: ISCED_7=code 3
Low: ISCED_7=code 1 or code 2

Level	Number
High	2552
medium	6562
low	3584
missing cases	160

EDU_Y: Year highest level of education achieved used: a150y

Missing cases_ 446

EDU_M: Month highest level of education achieved used: a150m

Missing cases: 887+seasonal codes

IEDU_Y: Year highest level education achieved and imputed year

Definition for imputation:

1) find the modal age of graduation (with help of graduation dates and birth dates for available cases) for every level of education. Year of graduation for missing cases then is calculated by adding modal age of graduation to the birth date (year and month).

After these imputations remain 12 unknown years

IEDU_M: Month highest education achieved and imputed month

Definition:
1) if only month unknown/ year known: find a random variable according to manual
2) if seasonal code – find a random variable according to manual
3) if month and year unknown use month achieved in process above (IEDU_Y)

After these imputations remain 9 unknown months

---

8. Part Background variables (ethnicity, nationality etc.)

NATIVE: Born in country used: a105

Born in country: 12725
Born elsewhere: 116
17 missing cases

ETHNOS: Ethnicity/nationality used: a110

Country specific variable (100+1+code)

missing cases: 57

BIRTH_COU: Country of birth used: a106b

Country specific variable (100+1+code)

Filter: BIRTH_COU=.b if a105==1

missing cases: 3

MIG_Y: Year of migration used: a107y

missing cases: 10

Filter: MIG_Y=.b if a105==1

MIG_M: Month of migration used: 107m

15 missing cases and additionally seasonal codes

Filter: MIG_M=.b if a105==1

IMIG_M: Month of migration and imputed months used: MIG_M

according to manual page 4 (random)

---

9. Part Background variables (parental background)

SIS_NO: Number of sisters used: a5106a_s

missing cases: 407

BRO_NO: Number of brothers used: a5106a_b

missing cases: 352

SIBS: Total number of sibs used: a5106a_s and a5106a_b
missing cases: 112

DECISION: If number of sisters is known and number of brothers is unknown or number of brothers is known and number of sisters is unknown: the number of known brothers or sisters is used

if number of brothers and number of sisters is unknown the value remains : missing (.a)

SIS_DIED: Number of sisters that died
used: a5106a_s and a5106b_s
(number of sisters respondent have ever had – number of alive sisters)

Filter: SIS_DIED=.b if a5106a_s==0
Missing cases: 444
Transformations: Negative values achieved➔ .a (missing)

BRO_DIED: Number of brothers that died
used: a5106a_b and a5106b_b

Filter: BRO_DIED=.b if a5106a_b==0
Missing cases: 380
Transformations: Negative value achieved➔ .a (missing)

ISCED_MO: Mother`s highest level of education used: a5115

ISCED	Number
0	616
1	1733
2	3861
3	3914
5	1198
6	9
.a	702
7	825

ISCED_FA: Father`s highest level of education used: a5113

ISCED	Number
0	515
1	1557

2	3988
3	3889
5	1066
6	18
.a	1239
7	586

EDU3_MO: Highest level of education of mother
ISCED 1997, collapsed into 3 categories used: ISCED_MO

Definition: 1 (high) if ISCED_MO=5 or 6
2 (medium) if ISCED_MO=3 or 4
3 (low) if ISCED_MO=1 or 2

Level	Number
High	1207
medium	7775
low	3174
missing cases	702

EDU3_FA: Highest level of education of father
ISCED 1997, collapsed into 3 categories used: ISCED_FA

Definition: 1 (high) if ISCED_FA=5 or 6
2 (medium) if ISCED_FA=3 or 4
3 (low) if ISCED_FA=1 or 2

Level	Number
High	1084
medium	7877
low	2143
missing cases	1239

WORK_MO: Mother`s occupation, when respondent was 15
Country codes used: 5114
Missing values: 198+2412

WORK_FA: Father`s occupation, when respondent was 15
Country codes used: 5112
WORK_FA missing cases: 729+830

ISCO3_MO: Mother`s occupation, when respondent was 15
3 categories used: WORK_MO

Definition: according to manual page 7
* Group 1: High non manual: 1, 2, 3
* Group 2: Non manual: 4, 5, 0
* Group 3: Manual: 6,7,8,9

Level	Number
1	2057
2	1959
3	6232
.a	2610

ISCO3_FA: Father`s occupation, when respondent was 15
3 categories used: WORK_FA

Definition: according to manual page 7
* Group 1: High non manual: 1, 2, 3
* Group 2: Non manual: 4, 5, 0
* Group 3: Manual: 6,7,8,9

Level	Number
1	2316
2	685
3	8298
.a	1559

NATIVE_MO: Mother born in country used: a513a
Mother born in country: 12542
Missing cases: 78
Born elsewhere: 238

NATIVE_FA: Father born in country used: a533a
Father born in country: 12476
Missing cases: 128
Born elsewhere: 254

BIRTHCO_MO: Mother`s country of origin, used: a513b

Country specific variable (100)

Filter: BIRTHCO_MO=.b if NATIVE_MO==1

BIRTHCO_MO missing cases: 90

BIRTHCO_FA: Father`s country of origin, used: a533b

Country specific variable (100)

Filter: BIRTHCO_FA=.b if NATIVE_FA==1

BIRTHCO_FA missing cases: 142

PARDIVEV: Parents ever divorced/separated

used: a516,a523,a542,a549,a571,a5104 (for every type of living with or without parents was asked one of these questions)

* a516 if respondent lives with biological father and mother dead/ or respondent do not know anything about mother
* a523 if respondent lives with biological father and mother alive
* a542 if respondent lives with biological mother and father dead/ or respondent do not know anything about father
* a549 if respondent lives with biological mother and father alive
* a571 if respondent lives without biological parents
* a5104 if respondent lives with both of his parents

Definition:

1) “Parents ever divorced/separated” (code 1) if: there is code 1 (yes, biological parents ever broke up) in the used questions 1068
2) “No-stayed together” (code 2) if: a5104==2 (they never broke up), or respondent lives without parent and they never separated (a571==2) and both are alive (a557 and a564==1) 10662
3) “They never lived together” (code 3) if: there is code 2 in the questions and code 3 in a571 349
4) “Parental death” (code 4) if: there is code 3 in q516 and q542 and a509/a535==2 or a571==2 & (a557==2 | a564==2) 571
5) “No, no other information available” (code 5) if: code 3 (no, another information) and no death 35

Missing cases: 173

Filter: .b if a509==4 | a535==4 | a564==4 (53 cases)

PARDIV_15: Parents divorced before age of 15

used: a516,a523,a542,a549,a571,a5104 (for every type of living with or without parents one of these questions was asked)

Definition:

1) “Parents divorced/separated” (code 1) before age 15 of respondent if: there is code 1 in the questions and year of separation-birth year of respondent <=15 (619 cases)
2) “No stayed together” (code 2)if respondent lives with both parents and they never separated or respondent lives without parents and they never separated and they are alive or other situation and mother or father were dead at the time of interview, but not at the age of 15 of respondent (11364 cases)
3) “They never lived together” (code 3) if there is code 2 in the questions or code 3 in q571 (349 cases)
4) “Parental death” (code 4) if: there is code 3 in the questions and mother or father died before age 15 of respondent (313 cases)
5) “no other information” (code 5) if: code 3 and no death (23 cases)

187 missing cases

---

10. Part Background variables (region, size of location)

REGION: Country region at time of interview
Country specific variable (100 +1 +code) used: aregion

No missing cases

SIZE: Size of place of residence at time of interview,

Country specific variable (100+1+code) used: atype

No missing cases

ISIZE: Size of place of residence at time of interview

Standardized code

SIZE_15: Size of place of residence at age 15 used: a5108

Missing cases: 125

ISIZE_15: Size of place of residence at age 15

Standardized code

---

11. Part Other background variables

RELIGION: Religious affiliation at time of interview
Country specific variable (100+1 +code) used: a1101

RELIGION=.a if a1101==97 | a1101==98

Missing cases: 35

IRELIGION: Religious affiliation at time of interview

Standardized code

ADOPT: Number of adopted children of respondent
used: ahg3_3-ahg3_6 (code5) and a213 (code 2)

FOSTER: Number of foster children of respondent
Used: ahg3_3-ahg3_6 (code 6) and a213 (code 3)

STEP: Number of stepchildren of respondent
Used: ahg3_3-ahg3_8 (code 4) and a226/ a229

Number of children	Adopt	Foster	Step
1	71	18	169
2	4	6	108
3		2	17
4			7
5			3
6			3
7			
8			

---

12. Part Weights

HHWGT: Household weight – not available in survey

PERSWGT: Personal weight - aweight

KISHWGT: Kishweight - not available in survey