

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

HARMONIZED HISTORIES Norway (14881 respondents)

Karolin Kubisch
Max Planck Institute for Demographic Research Rostock

2010
Updated 3.6.2014
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: UN Data: GGS_Wave1_Norway_V3.0.dta

Interview dates Norwegian GGS: 2007/2008

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 Norway it affects ca. 125 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 marriage date and negative difference between divorce date and union or marriage date
- Sucessive partnerships $\text{mar}-\text{mar}[_n-1] \leq 0$ or $\text{par}-\text{par}[_n-1] \leq 0$
- Differences between separation date and next partnership date $\text{sep} > \text{par}[_n+1]$

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) 14881 respondents	used: arid
COUNTRY:	Country and survey COUNTRY: code: 5781: Norway GGS	used: acountry
MONTH_S:	Month of survey	used: amonth
IMONTH_S:	Month of survey, including imputed dates	
YEAR_S:	Year of survey 2007/2008	used: ayear
SEX:	Sex of the respondent Sex structure of the Norwegian respondents: Male: 7340 and Female: 7541	used: ahg4_1
BORN_Y:	Year of birth of respondent 1927-1988	used: ahg6y_1
BORN_M:	Month of birth of respondent	used: ahg6m_1
IBORN_M:	Month of birth of respondent including imputed months Harmonized: random variable between 1-12	used: BORN_M

2. Part LEAVING HOME

LEAVE_1: Indicator of whether "left home"
used: GRID=1 go to a5117a
=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: 361 / 1: 14512
8 missing cases

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

Filter: LEAVE_Y1/LEAVE_M1 to .b if LEAVE_1==0 (361)
Missing cases: 557

LEAVE_M1: Month of first time leaving home used: a5116m and a5117bm

Missing cases: .b 361 .a 1865

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

Harmonized: random variables according to manual

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

UNINUM: Total number of unions used: UNION_1 to _5

Syntax:
forvalues x=1/5 {
replace UNINUM=UNINUM+1 if UNION_`x'>0
}

UNINUM:
0: 2169
1: 9453
2: 2749
3: 438
4: 62
5: 9

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):

→an 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: 12711
UNION_2: 3258
UNION_3: 509
UNION_4: 71

UNION_5: 9

Transformations

```
replace a301y=.a if ARID==19664 | ARID==1225 | ARID==1490 | ARID==10774
| ARID==11754 | ARID==12055 | ARID==16082 | ARID==16417 | ARID==18593 |
ARID==20217 | ARID==41194 | ARID==549
replace a302by=.a if ARID==41846
replace a301y=.a if ARID==1314 | ARID==1588 | ARID==2571 | ARID==2687
| ARID==3313 | ARID==3592 | ARID==4092 | ARID==16553 | ARID==18472 |
ARID==18918 | ARID==21486 | ARID==40079 | ARID==41092
replace a301m=11 if ARID==1588 | ARID==3828 | ARID==16470 |
ARID==18875 | ARID==21901 | ARID==42564 | ARID==42274 | ARID==20915
replace a301y=.a if ARID==5054 | ARID==8008 | ARID==8021 | ARID==9000 |
ARID==10457 | ARID==11375 | ARID==12492 | ARID==15208 | ARID==13045
replace a301m=8 if ARID==5838 | ARID==15536 | ARID==17393 | ARID==40478
replace a301m=12 if ARID==7077
replace a301m=5 if ARID==8747 | ARID==40854
replace a334y_1=.a if ARID==21045 | ARID==21394 | ARID==41367 |
ARID==340
replace a335y_1=.a if ARID==2593 | ARID==9545 | ARID==544 | ARID==3304
| ARID==41685
replace a344y_1=.a if ARID==559 | ARID==1056 | ARID==1621 | ARID==2274
| ARID==3150 | ARID==8580 | ARID==9709 | ARID==11645 | ARID==13204 |
ARID==16284 | ARID==16303 | ARID==16435 | ARID==17279 | ARID==17461 |
ARID==18594 | ARID==19404 | ARID==41829 | ARID==6244 | ARID==8937 |
ARID==549 | ARID==4303 | ARID==20195
replace a344y_2=.a if ARID==1562 | ARID==1680 | ARID==3560 |
ARID==8937 | ARID==10170 | ARID==16650 | ARID==20195 | ARID==1744 |
ARID==5834 | ARID==5980 | ARID==7978 | ARID==12327 | ARID==12694 |
ARID==13377 | ARID==15141 | ARID==16206 | ARID==17144 | ARID==17159 |
ARID==20223 | ARID==20697 | ARID==21686 | ARID==41151
replace a335y_2=.a if ARID==7460 | ARID==1694 | ARID==7556 | ARID==1744
| ARID==5834 | ARID==18850
replace a334y_2=.a if ARID==10170 | ARID==17144 | ARID==17322 |
ARID==18850 | ARID==21686 | ARID==40013 | ARID==41151 | ARID==41538
replace a334m_2=11 if ARID==23
replace a334m_3=8 if ARID==240
replace a334m_2=8 if ARID==403 | ARID==42333
replace a334m_2=11 if ARID==868
replace a334y_2=.a if ARID==2789 | ARID==2807 | ARID==4151 | ARID==9067
| ARID==11587 | ARID==12327 | ARID==13566 | ARID==16206
replace a334m_2=11 if ARID==9884 | ARID==42402
replace a334m_3=8 if ARID==9884
replace a334m_2=11 if ARID==13045 | ARID==19704
replace a334m_2=9 if ARID==14572 | ARID==10586
replace a334m_2=5 if ARID==19582
```

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

Filter: UNION_Yx=.b if UNION_x==0

UNION_Y1 missing values: 569
UNION_Y2 missing values: 143
UNION_Y3 missing values: 24
UNION_Y4 missing values: 7

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

Filter: UNION_Mx=.b if UNION_x==0

UNION_M1 missing values: 1764
UNION_M2 missing values: 433
UNION_M3 missing values: 56
UNION_M4 missing values: 12

IUNION_M\$: Month of start UNION used: UNION_M\$
and imputed months

according to manual page 4 (random)

Filter: IUNION_Mx=.b if UNION_x==0

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

Filter: SEP_x=.b if UNION_x==0
in case of current partner: no separation

Order of Union	Number of unions	number of separations	death of partner	UNK
1	12711	4492	623	25
2	3258	967	71	10
3	509	163	9	2
4	71	26		2
5	9	3		

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

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

SEP_Y1 missing values: 194
SEP_Y2 missing values: 78
SEP_Y3 missing values: 18
SEP_Y4 missing values: 4

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: 1006
SEP_M2 missing values: 232
SEP_M3 missing values: 34
SEP_M4 missing values: 7

ISEP_M\$: Month of end of UNION used: SEP_M\$
and imputed months

according to manual page 4 (random)

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

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: 13
MARR_2 missing values: 9
MARR_3 missing values: 2
MARR_4 missing values: 1

Order of Union	Number of unions	number of marriages
1	12711	9036
2	3258	1616
3	509	215
4	71	27
5	9	4

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: 527
MARR_Y2 missing values: 60
MARR_Y3 missing values: 11
MARR_Y4 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: 798
MARR_M2 missing values: 104
MARR_M3 missing values: 14
MARR_M4 missing values: 2

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

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

DIV_\$: Indicator of whether divorce occurred

NOT INCLUDED IN SURVEY

DIV_Y\$: Year of divorce

NOT INCLUDED IN SURVEY

DIV_M\$: Month of divorce

NOT INCLUDED IN SURVEY

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

NOT INCLUDED IN SURVEY

Summary:

The variables DIV_\$, DIV_Y\$, DIV_M\$ and IDIV_M\$ are not included in dataset.

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

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

Filter: SEXP_x=.b if UNION_x==0

Partner	Number of unions	Number male	Number female
1	12711	6663	6048
2	3258	1734	1524
3	509	287	222
4	71	37	34
5	9	3	6

YEARBIRP_\$: Year of birth of partner Used: ahg6y_2

a336y → not included in survey → .c (only for current partner)

Filter: YEARBIRP_x=.b if UNION_x==0

YEARBIRP_1 missing cases: 19
YEARBIRP_2 missing cases: 17
YEARBIRP_3 missing cases: 2
YEARBIRP_4 missing cases: 2

MONBIRP_\$: Month of birth of partner used: ahg6m_2

a336m → not included in survey → .c (only for current partner)

Filter: MONBIRP_x=.b if UNION_x==0

MONBIRP_1 missing cases: 19
MONBIRP_2 missing cases: 17
MONBIRP_3 missing cases: 2

MONBIRP_4 missing cases: 2

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

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 : code 4: stepchild: my current partners child not
adopted by me → ahg3_

b)non-resident stepchildren: a226==1 and a229

c)for partnership histories: a338_

also: year of start of union(a301y) and year of birth of stepchild
(ahg6y_x and a230_x)

Problem: The question: When you started living together, how many
children did your partner have? (a338)- exists only for partnership
histories

-for current partnership it had to be created with help of the number
of stepchildren, year of start of union and year of birth of stepchild

Definition: in the number of children of current partner are included:

* all stepchildren of respondent living at the moment of interview in
household grid and were born before the start of the union

* all nonresident stepchildren at the time of interview - partners
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_)

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

NUMCHP_1: missing values: 74

NUMCHP_2: missing values: 6

NUMCHP_3: missing values: 2

NUMCLIV_ \$: Number of children of partner lived with respondent

a341_1 - a341_6 not included in survey

Union	Number of unions	NUMCHP
1	12711	1-7,13
2	3258	1-6
3	509	1-5,8
4	71	1-6
5	9	1-3

Summary:

The variable NUMCLIV is not included in dataset.

6. Part Birth histories (biological kids)

For the chapter "Birth histories" an reshaping program was used, which includes biological children in household and questions 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 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
1	10718
2	8717
3	4018
4	1087
5	267
6	83
7	36
8	14
9	6
10	3
11	2

KID_Y\$: Year of birth of child

used: ahg6y_ and a216y

Filter: KID_Yx=.b if KID_x==0

KID_Y1 missing values: 14

KID_Y2 missing values: 11

KID_Y3 missing values: 12

KID_Y4 missing values: 5

KID_Y5 missing values: 2

KID_Y6 missing values: 2

KID_Y7 missing values: 2

TRANSFORMATIONS

replace birth_r=.a if ARID==375

replace birth_r=1957 if ARID==4652

replace birth_r=.a if ARID==5310

replace birth_r=.a if ARID==9361

replace birth_r=.a if ARID==41183

replace a220y_4=.a if ARID==1148

replace a220y_1=.a if ARID==1178

8	14	3	11
9	6	4	2
10	3	2	1
11	2		2

KID_D\$: Death of child used: a211b

Filter: KID_Dx=.b if KID_x==0

No death children → there are no cases for biological children where a211b=yes. (all these cases (death) are coded in a213_ with not asked)

KID_DY\$: Year of death of child used: a217y

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

NOT INCLUDED IN SURVEY

KID_DM\$: Month of death of child used: a217m

NOT INCLUDED IN SURVEY

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

NOT INCLUDED IN SURVEY

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	10718	5938
2	8717	4548
3	4018	2012
4	1087	585

5	267	169
6	83	60
7	36	21
8	14	8
9	6	4
10	3	1
11	2	

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: 354
KID_LY2 missing cases: 254
KID_LY3 missing cases: 152
KID_LY4 missing cases: 57
KID_LY5 missing cases: 25
KID_LY6 missing cases: 15
KID_LY7 missing cases: 5
KID_LY9 missing cases: 1
KID_LY10 missing cases: 1

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: 1468
KID_LM2 missing cases: 1178
KID_LM3 missing cases: 591
KID_LM4 missing cases: 200
KID_LM5 missing cases: 74
KID_LM6 missing cases: 33
KID_LM7 missing cases: 12
KID_LM8 missing cases: 5
KID_LM9 missing cases: 1
KID_LM10 missing cases: 1

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

Summary: No information about death of children are included in survey.

7. Part Education

INSCHOOL: Currently studying at the time of interview

used: a151

Currently studying: 995
Missing cases: 85

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

These data exist in the harmonized dataset in an ISCED97 coded form.

These country specific codes include:

- * a 3-digit country prefix(578)
- * a 1-digit survey code (Norway GGS=1) and
- * a 2-digit country specific code for level of education (1-7 levels of education)

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

Harmonized:

ISCED	Number
0+1	24
2	2994
3	6354
4	475
5	4881
6	79
.a	174

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

Level	Number
High	4960
medium	6729
low	3018
.a	174

EDU_Y: Year highest level of education achieved used: a150y

Missing cases: 146

Special code for NORWAY: Norway deviation 1970 and earlier (.d)

EDU_M: Month highest level of education achieved

Missing cases: 146

Special code for NORWAY: Norway deviation 1970 and earlier (.d)

IEDU_Y: Year highest level education achieved and imputed year

NO change of unknown cases, because of the special Norwegian code

IEDU_M: Month highest education achieved and imputed month
replace IEDU_M=int(uniform()*12)+1 if IEDU_M=.

Summary:

The EDU_COU data exist in a country specific ISCED97 form.

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

NATIVE: Born in country used: a105

Born in country: 13940
Born elsewhere: 941

ETHNOS: Ethnicity/nationality

NOT INCLUDED IN SURVEY

BIRTH_COU: Country of birth used: a106b

Country specific variable (578+1+code)

Filter: BIRTH_COU=.b if a105==1
Missing cases: 1

MIG_Y: Year of migration used: a107y

Missing cases: 2

Filter: MIG_Y=.b if a105==1

MIG_M: Month of migration used: 107m

Missing cases: 45 and additional 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)

Summary:

The variable ETHNOS is not included in survey.

9. Part Background variables (parental background)

SIS_NO: Number of sisters used: a5106a_s

0 - 10 sisters

missing cases: 1026

BRO_NO: Number of brothers used: a5106a_b

0 - 10 brothers

missing cases: 1028

SIBS: Total number of sibs used: a5106a_s and a5106a_b

0-18 sibs

missing cases: 1027

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: 1029

BRO_DIED: Number of brothers that died

used: a5106a_b and a5106b_b

Filter: BRO_DIED=.b if a5106a_b==0

Missing cases: 1031

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

NOT INCLUDED IN SURVEY

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

NOT INCLUDED IN SURVEY

EDU3_MO: Highest level of education of mother

ISCED 1997, collapsed into 3 categories

used: ISCED_MO

Level	Number
High	2002
medium	5277
low	6792
.a	1859

EDU3_FA: Highest level of education of father

ISCED 1997, collapsed into 3 categories

used: ISCED_FA

Level	Number
High	2424

medium	5431
low	5167
.a	1859

WORK_MO: Mother`s occupation, when respondent was 15
Country codes used: 5114
Missing values: 6660

WORK_FA: Father`s occupation, when respondent was 15
Country codes used: 5112
Missing values: 2747

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

Filter: ISCO3_MO=.b if WORK_MO=.b

Level	Number
1	2224
2	3908
3	2089
.a	6660

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	2853
2	2395
3	6886
.a	2747

NATIVE_MO: Mother born in country used: 513a

Missing cases: 14030

NATIVE_FA: Father born in country used: 533a

Missing cases: 14211

BIRTHCO_MO: Mother`s country of origin used: a513b

country specific variable

Filter: BIRTHCO_MO=.b if NATIVE_MO==1

BIRTHCO_MO missing cases: 14795

BIRTHCO_FA: Father`s country of origin used: a533b

country specific variable

Filter: BIRTHCO_FA=.b if NATIVE_FA==1

missing cases: 14804

PARDIVEV: Parents ever divorced/separated used: a550/a552

Missing values: 547

PARDIV_15: Parents divorced before age of 15 used: a550/a552
a551/ a511/ ahg6y_1

missing values: 3566

Background variables (region, size of location)

REGION: Country region at time of interview

Country specific variable used: aregion

missing cases:1

SIZE: Size of place of residence at time of interview used: atype

Country specific variable

missing cases: 1

ISIZE: Size of place of residence at time
of interview

Standardized code

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

NOT INCLUDED IN SURVEY

ISIZE_15: Size of place of residence at age 15

Standardized code

Summary:

The variable SIZE_15 is not included in survey.

11. Part Other background variables

RELIGION: Religious affiliation at time of interview

Country specific variable

used: a1101

Missing values: 133

IRELIGION: Religious affiliation at time of interview

Standardized code

ADOPT: Number of adopted children of respondent

used: ahg3_2-ahg3_5, ahg3_8 (code5) and a213 (code 2)

NO adopted children

FOSTER: Number of foster children of respondent

Used: ahg3_2-ahg3_6 (code 6) and a213 (code 3)

STEP: Number of stepchildren of respondent

Used: ahg3_2-ahg3_8 (code 4) and a226/ a229

Number of children	Foster	Step
1	56	633
2	16	525
3	4	181
4		48
5		5
6		3

12. Part Weights

HHWGT: Household weight - not available in survey

PERSWGT: Personal weight - aweight

KISHWGT: Kishweight - not available in survey