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

HARMONIZED HISTORIES Italy (9570 respondents age 18 to age 64)

Karolin Kubisch
Max Planck Institute for Demographic Research Rostock

2012
Updated 3.6.2014
Updated 27.10.2015
Update 11.04.2019

```
The following document 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.

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

Source: GGS First Wave, UN version 2019: GGS_Wave1_Italy_V.4.3.dta

Interview dates Italy GGS: 2003

```
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 Italy it affects 0 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

1.Part Basic Information

RESPID: ID number to be assigned at merging   LEAVE BLANK

ARID: ID number from raw data (original ID number)
              9570 respondents

COUNTRY: Country and survey
              Harmonized: code: 3801: Italy GGS

MONTH_S: Month of survey
              Missing cases: 492

IMONTH_S: Month of survey, including imputed dates

YEAR_S: Year of survey
              2003

SEX: Sex of the respondent
              No missing cases
              Sex structure of the Italian respondents:
              Female: 5115
              Male: 4455

BORN_Y: Year of birth of respondent
              UNAVAILABLE IN SURVEY

BORN_M: Month of birth of respondent
              UNAVAILABLE IN SURVEY

IBORN_M: Month of birth of respondent
              including imputed months
              UNAVAILABLE IN SURVEY

AGE_R: Age of respondent (generated only for Italy because of
              missing month and year of birth)
              18-64 years

Summary:
BORN_Y and BORN_M are not available. That's why I generated a variable AGE_R (age of respondent) from the UN version for Italy.

2. Part LEAVING HOME

LEAVE_1: Indicator of whether left home
0 (No): 2029 / 1 (yes): 7541

LEAVE_Y1: Year of first time leaving home

Filter: .b if LEAVE_1==0 (2029)

LEAVE_M1: Month of first time leaving home

UNAVAILABLE IN SURVEY

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

UNAVAILABLE IN SURVEY

Summary: The month of leaving home is not included in
survey.

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
          }
```

0: 2600
1: 6681
2: 261
3: 24
4: 4

UNION_$: UNION order

UNION_1: 6970
UNION_2: 289
UNION_3: 28
UNION_4: 4

No missing cases

UNION_Y$: Year of start union

UNION_Yx=.b if UNION_x==0

UNION_Y1 missing values: 149
UNION_Y2 missing values: 7

TRANSFORMATIONS:
replace a301y=.a if ARID==10930 | ARID==25510 | ARID==30296 |
ARID==32359 | ARID==41667 | ARID==48071 | ARID==54261 | ARID==59442 |
ARID==60138 | ARID==71186 | ARID==71678 | ARID==74555
replace a301y=1998 if ARID==21374
replace a301y=2002 if ARID==22197
replace a301y=1991 if ARID==29031
replace a301y=2001 if ARID==33964
replace a301y=.a if ARID==61032 | ARID==64909 | ARID==19713 |
ARID==83385 | ARID==87833
replace a344y_1=1988 if ARID==33964

UNION_M$: Month of start UNION

Filter: UNION_Mx=.b if UNION_x==0

UNION_M1 missing values: 153
UNION_M2 missing values: 6
UNION_M3 missing values: 2

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

Filter: IUNION_Mx=.b if UNION_x==0

SEP_$:   Dissolution of UNION

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	missing
1	6970	570	327	272
2	290	20	7	67
3	28	2		9
4	4			1

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: 173
SEP_Y2 missing values: 8
SEP_Y3 missing values: 2

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: 952
SEP_M2 missing values: 33
SEP_M3 missing values: 3

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

Filter: MARR_x=.b if UNION_x==0

Order of Union	Number of unions	number of marriages
1	6970	6647
2	290	148
3	28	13
4	4	3

MARR_Y$:   Year of marriage

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

MARR_M$: Month of marriage

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

IMARR_M$: Month of marriage
               and imputed months
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

Filter: DIV_x=.b if UNION_x==0

```
                DIV_x=.b if MARR_x==0
```

Order of Union	Number of unions	number of marriages	number of divorces
1	6970	6647	263
2	290	148	6
3	28	13	1
4	4	3	

DIV_1 missing values: 55
DIV_2 missing values: 6
DIV_3 missing values: 1

DIV_Y$:   Year of divorce

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: 55
DIV_Y2 missing values: 6
DIV_Y3 missing values: 1

DIV_M$:   Month of divorce
               UNAVAILABLE IN SURVEY

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

               UNAVAILABLE IN SURVEY

Summary: The month of divorce is not included in survey.

---

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

SEXP_$:   Partner`s sex

Filter: SEXP_x=.b if UNION_x==0

YEARBIRP_$:   Year of birth of partner

               UNAVAILABLE IN SURVEY

MONBIRP_$:   Month of birth of partner

               UNAVAILABLE IN SURVEY

IMONBIRP_$:   Month of birth of partner
                   and imputed months

               UNAVAILABLE IN SURVEY

6

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

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

               UNAVAILABLE IN SURVEY

NUMCLIV_$:   Number of children of partner lived with respondent

               UNAVAILABLE IN SURVEY

Summary: The variables YEARBIRP_, MONBIRP_ NUMCLIV_ and
NUMCHP_ are not included

---

6. Part Birth histories (biological kids)

KID_$: Indicator of child order

no missing cases

Child order	number of children
1	6124
2	4191
3	1066
4	225
5	46
6	11
7	6
8	2
9	1

Problems with birth intervals <7 months or >20 years:
(no changes, only for information)

```
ARID   SEX             Births(M,Y)
25709 Male             7/67 – 12/68 – 10/69 – 3/70        possible
25824 Female           9/66 – 5/67 - 11/70
26631 Male             8/71 – 12/73 – 1/74                possible
27007 Female           9/61 – 5/62 - 10/64
36877 Female           12/63 – 1/64
38884 Female           12/67 – 1/68
64732 Male             12/76 – 1/77                       possible
67328 Female           9/69 -12/71- 1/72
71083 Female           11/58 – 10/62 – 1/63
71948 Female           12/69 – 5/70
83915 Female           5/61 – 10/63 – 1/64
87933 Female           11/63 – 1/64
```

KID_Y$: Year of birth of child (not available for hh members)

Filter: KID_Yx=.b if KID_x==0

KID_Y for Kids in household is not available from survey
In 2019 the data for them were constructed from year of survey and age
at interview
rename ahg5_ KID_AGE
g KID_Y=year_s-KID_AGE
g KID_M=.a if !missing(KID_Y)

KID_M$: Month of birth of child (not available for hh members)

Filter: KID_Mx=.b if KID_x==0

(KID_M for Kids in household is not available)
KID_M1 missing values (.c): 4008
KID_M2 missing values: 2986
KID_M3 missing values: 707
KID_M4 missing values: 129
KID_M5 missing values: 22
KID_M6 missing values: 4
KID_M7 missing values: 2
KID_M8 missing values: 2
KID_M9 missing values: 1

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

Filter: IKID_M_x=.b if KID_x==0

KID_S$: Sex of child

Filter: KID_Sx=.b if KID_x==0

Child order	number of children	male	female
1	6124	3193	2931
2	4191	2159	2032
3	1066	522	544
4	225	115	110
5	46	20	26
6	11	6	5
7	6	3	3
8	2	1	1
9	1	1	

KID_D$: Death of child

Filter: KID_Dx=.b if KID_x==0

No missing cases

Child order	number of children	death
1	6124	70

2	4191	45
3	1066	20
4	225	4
5	46	2
6	11	
7	6	
8	2	
9	1	

KID_DY$: Year of death of child

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

KID_DM$: Month of death of child

   NOT INCLUDED IN SURVEY

IKID_DM$: Month of death of child
             and imputed months

   NOT INCLUDED IN SURVEY

KID_L$:   Child left home

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.

Filter: KID_Lx=.b if KID_x==0

Child order	number of children	Left home
1	6124	2046
2	4191	1160
3	1066	339
4	225	92
5	46	22
6	11	7
7	6	4
8	2	
9	1	

KID_LY$: Year child left home

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

KID_LM$: Month child left home

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

        NOT INCLUDED IN SURVEY

IKID_LM$: Month of death of child
            and imputed months

according to manual page 4 (random variable)

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

        NOT INCLUDED IN SURVEY

Summary:
The variables KID_DM, IKID_DM, KID_LM, IKID_LM are not
included. Problematical cases in birth intervals between 2
births were not changed.

7. Part Education

INSCHOOL: Currently studying at the time of interview

        NOT INCLUDED IN SURVEY

EDU_COU: Highest level of education, country specific

the country specific codes include:
* a 3-digit country prefix(380)
* a 1-digit survey code (Italy GGS=1) and
* a 2-digit country specific code for level of education (1-9 levels of
education)

ISCED_7: Highest level of education
          Achieved according to ISCED 1997

ISCED	Number
0+1	1779
2	2888
3	3861
10 (5A-6)	57
11 (5A)	900
12 (5A-5B)	85

EDU_3:   Highest level of education ISCED

```
          Collapsed into 3 categories
```

Definition: High: 5+6
          Medium: 3
          Low: 0+1+2

Level	Number
High	1042
medium	3861
low	4667

EDU_Y: Year highest level of education achieved

Filter: .b 168

EDU_M: Month highest level of education achieved

    NOT INCLUDED IN SURVEY

IEDU_Y: Year highest level education achieved and imputed year

    NOT INCLUDED IN SURVEY

IEDU_M: Month highest education achieved and imputed month

    NOT INCLUDED IN SURVEY

Summary: The variable EDU_M is not included in survey. That's why also IEDU_Y and IEDU_M were not generated. There are 3 country specific ISCED codes for Italy.

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

NATIVE: Born in country

Born in country: 9070
Born elsewhere: 275

Missing values: 225

ETHNOS: Ethnicity/nationality

    NOT INCLUDED IN SURVEY

BIRTH_COU: Country of birth

    NOT INCLUDED IN SURVEY

MIG_Y: Year of migration

MIG_M: Month of migration

IMIG_M: Month of migration and imputed months

Summary: The variables ETHNOS, BIRTH_COU, MIG_Y, MIG_M, IMIG_M are not included.

---

9. Part Background variables (parental background)

SIS_NO: Number of sisters

0 – 13 sisters

BRO_NO: Number of brothers

0 – 11 brothers

SIBS: Total number of sibs

0-15 sibs

SIS_DIED: Number of sisters that died

BRO_DIED: Number of brothers that died

ISCED_MO: Mother`s highest level of education

ISCED	Number
0+1	6626
2	1416
3	1079
10	14
11	178
12	25
.a	232

ISCED_FA: Father`s highest level of education

ISCED	Number
0+1	6143

12

2	1573
3	1198
10	45
11	297
12	34
.a	280

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

Definition: 1 (high) if code 5+6
             2 (medium) if code 3
             3 (low) if code 0+1+2

Level	Number
High	217
medium	1079
low	8042
.a	232

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

Definition: 1 (high) if code 5+6
             2 (medium) if code 3
             3 (low) if code 0+1+2

Level	Number
High	376
medium	1198
low	7716
.a	280

WORK_MO: Mother`s occupation, when respondent was 15
Country codes
Missing values: 109

WORK_FA: Father`s occupation, when respondent was 15
Country codes
Missing values: 160

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

Definition: according to manual page 7
* Group 1: High non manual
* Group 2: Non manual
* Group 3: Manual

        NOT INCLUDED IN SURVEY

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

13

Definition: according to manual page 7
* Group 1: High non manual
* Group 2: Non manual
* Group 3: Manual

NOT INCLUDED IN SURVEY

NATIVE_MO: Mother born in country

NOT INCLUDED IN SURVEY

NATIVE_FA: Father born in country

NOT INCLUDED IN SURVEY

BIRTHCO_MO: Mother`s country of origin, country specific (380)

NOT INCLUDED IN SURVEY

BIRTHCO_FA: Father`s country of origin, country specific (380)

NOT INCLUDED IN SURVEY

PARDIVEV: Parents ever divorced/separated

Missing values: 2

PARDIV_15: Parents divorced before age of 15

NOT INCLUDED IN SURVEY

Summary: The variables SIS_DIED, BRO_DIED, ISCO3_MO, ISCO3_FA, NATIVE_MO, NATIVE_FA, BIRTHCO_MO, BIRTHCO_FA, PARDIV_15 are not included.

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

REGION: Country region at time of interview

Country specific variable (380 +1 +code)

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

NOT INCLUDED IN SURVEY

ISIZE: Size of place of residence at time
               of interview

NOT INCLUDED IN SURVEY

SIZE_15: Size of place of residence at age 15

NOT INCLUDED IN SURVEY

ISIZE_15: Size of place of residence at age 15

NOT INCLUDED IN SURVEY

Summary: The variables SIZE and SIZE_15 are not included.

---

11. Part Other background variables

RELIGION: Religious affiliation at time of interview
            Country specific (250+1 +code)

NOT INCLUDED IN SURVEY

IRELIGION: Religious affiliation at time of interview

NOT INCLUDED IN SURVEY

ADOPT: Number of adopted children of respondent

NOT INCLUDED IN SURVEY

FOSTER: Number of foster children of respondent

NOT INCLUDED IN SURVEY

STEP: Number of stepchildren of respondent

Number of children	Adopt
1	64
2	44
3	11
4	4
5	1
7	1

Summary: The variables FOSTER and ADOPT are not included.

---

12. Part Weights
HHWGT: Household weight – NOT INCLUDED IN SURVEY

PERSWGT: Personal weight

KISHWGT: Kishweight   - NOT INCLUDED IN SURVEY