SOEP

The German Socio-Economic Panel study

580

SOEP Survey Papers

Series D – Variable Descriptions and Coding

SOEP-Core v33.1 – BIOAGEL: Generated variables from the "Mother & Child", "Parent", and "Pupils" questionnaires

Valeriia Heidemann, David Richter, and SOEP Group

DIW SOEP

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**

Please cite this paper as follows:

Valeriia Heidemann, David Richter, and SOEP Group. 2018. SOEP-Core v33.1 – BIOAGEL: Generated variables from the "Mother & Child", "Parent", and "Pupils" questionnaires. SOEP Survey Papers 580: Series D. Berlin: DIW/SOEP

SOEP-Core v33.1 – BIOAGEL: Generated variables from the "Mother & Child", "Parent", and "Pupils" questionnaires

Valeriia Heidemann, David Richter, and SOEP Group

BIOAGEL: Generated variables from the "Mother & Child", "Parent", and "Pupils" questionnaires

by Valeriia Heidemann & David Richter[1]

1   Introduction

The bioage data files are generated using information collected in the "Mother & Child" and "Parent" questionnaires. In 2014, for the first time, the children themselves answered the "Pupils" questionnaire. In addition, the data from the Families in Germany (FiD) study was integrated into the SOEP in 2014 as well. There are six different "Mother & Child" and "Parent" questionnaires and two "Pupils" questionnaires (see Table 1). Two additional questionnaires (biage02 and bioage10b) were only surveyed in FiD but are integrated into the *bioagel* data set, too.

Table 1:   **Overview of bioage data, corresponding age group, and respondents**

Bioage	Classification by Age	Respondents	First Wave
Bioage01	0-1 years old	Mothers only	2003
Bioage02 (FiD only)	1-2years old	Mothers only	2010
Bioage03	2-3 years old	Mothers only	2005
Bioage06	5-6 years old	Mothers only	2008
Bioage08a	7-8 years old	Parent 1	2010
Bioage08b	7-8 years old	Parent 2	2010
Bioage10	9-10 years old	Parent 1 (SOEP & FiD)	2012 / 2010 (FiD)
Bioage10b (FiD only)	9-10 years old	Parent 2 (FiD only)	2010
Bioage12	11-12 years old	Children	2014
Bioage14	13-14 years old	Children	2016

The "Mother & Child", "Parent" and "Pupils" questionnaires aim to follow and observe future generations of the SOEP and collect all information in age-specific files, even though the data come from different survey years. As we try to make this process as comprehensive and gap-free as possible, we begin following and documenting the development of children in SOEP households from birth onwards. Since many questions overlap, all bioage data files are covered in this chapter, rather than dedicating a separate chapter of this document to each bioage data

---

[1] This documentation is based on earlier versions of documentation materials on bioage data sets and has benefited from previous work by Sebastian Frischholz, Anne Fromm, Stefanie Lenuweit, Katharina Mahne, Christian Schmitt, and Jürgen Schupp.

file. By doing so, we are able to provide you with an overview of all variables covered in the bioage data files.

Starting with the v31 data distribution, all available data is now provided in one data set: *bioagel*. The separate bioage data files are no longer part of the data distribution.

2 Respondents in the 'Bioagel' Data Set

The values of the BIOAGE variable in the *bioagel* data file reflects the age of the children when the respective questionnaire was taken by their parents (e.g., *bioage* = 6 covers children who turned six that survey year, producing a range in ages between 5 years and 1 month and 6 years and 11 months, depending on the birth month and interview month). For information on the questionnaires used with each age group, please see Table 1. Except for bioage08b, bioage12, and bioage14 (and – FiD only – bioage10b), it is usually the mother who completes the questionnaire. In the exceptional case that a mother cannot complete the questionnaire, the father does. The respondents' personal ID numbers (regardless if fathers, mothers or other care takers) are provided in the variable PERSNRESP in the *bioagel* data set. Note that while the parents are the actual respondents (except for bioage12 & bioage14, covering the "Pupils" questionnaire), the data are organized under the child's unchanging personal ID number (PERSNR).

Bioage01: "Mother and Child" Questionnaire, Children Aged 0-1 years

The questionnaire is given to all women who gave birth to a child in the current or previous survey year, and to all women whose non-biological child was born in the same period. The questionnaire contains information on pregnancy, childbirth, childcare, and child health. In the exceptional case that a mother is unable to complete the questionnaire, the father responds.

Bioage02: "Your child between the ages of one and two", Children Aged 1-2 years (FiD only)

The questionnaire is given to all mothers whose child turns two in the current survey year. In the exceptional case that a mother cannot complete the questionnaire, the father responds.

Bioage03: "Your child between the ages of two and three", Children Aged 2-3 years

The questionnaire is given to all mothers whose child turns three in the current survey year. In the exceptional case that a mother cannot complete the questionnaire, the father responds.

Bioage06: "Your child between the ages of five and six", Children Aged 5-6 years

The questionnaire is given to all mothers whose child turns six in the current survey year. In the exceptional case that a mother cannot complete the questionnaire, the father responds.

Bioage08a and Bioage08b: "Parent" Questionnaire, children aged 7-8 years

The questionnaire is given to *both* mothers and fathers of children turning eight in the current survey year. Data of parent 1, which is usually the mother, can be found in *bioage08a*. Data of parent 2, which is usually the father, can be found in *bioage08b*.

Bioage10: "Your child between the ages of nine and ten", Children Aged 9-10 years (only FiD: Bioage10b)

The questionnaire is given to all mothers (FiD: and fathers) whose child turns ten in the current survey year. In the exceptional case that a mother cannot complete the questionnaire, the father responds. Data of parent 1, which is usually the mother, can be found in *bioage10a*. Data of parent 2 (FiD only), which is usually the father, can be found in *bioage10b*.

Bioage12: "Pupils between the ages of eleven and twelve", Children Aged 11-12 years

The questionnaire is given to all children themselves who turn(ed) twelve in the survey year.

Bioage14: "Pupils between the ages of thirteen and fourteen", Children Aged 13-14 years

The questionnaire is given to all children themselves who turn(ed) fourteen in the survey year.

Number of Children and Twins in the *Bioage* Data Sets

The data set has grown over the years and now contains data on 13,363 children. At the moment (2016), there are 66 children for whom information has been provided through all of the "Mother & Child", "Parent", and "Pupils" questionnaires (*bioage01*, *bioage03*, *bioage06*, *bioage08a/b, bioage10, bioage12, and bioage14*). Data are available for 170 children from seven questionnaires, 339 children from six questionnaires, 1,044 children from five questionnaires, 2,149 children from four questionnaires, for 2,031 from three questionnaires, and for 3,269 children from two questionnaires. For 4,295 children, data are available from one questionnaire. An overview of the number of children in each data set is given in Table 3.

Table 2: Number of respondents to the "Mother & Child", "Parent", and "Pupils" questionnaires

Bioage	Survey Year														Total
	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	
$01	318	247	246	234	205	185	196	1503	353	378	349	444	302	349	5309
$02 (FiD)								787	647	568	187				2189
$03			257	222	237	246	186	1061	914	745	701	454	386	395	5804
$06						237	210	687	696	612	858	825	657	686	5468
$08a								646	703	715	647	629	779	716	4835
$08b								409	490	500	439	399	556	535	3328
$10								403	510	699	686	656	616	539	4109
$10b (FiD)								242	310	291	301				1144
§12												606	608	563	1777
§14														531	531
Total	318	247	503	456	442	668	592	5738	4623	4508	4168	4013	3904	4314	34494

Table 4 gives an overview of the number of children per family in the specific age range. Note that there might be more children in the household than indicated here: these children are not in the age range of the "Mother & Child", "Parent", and "Pupils" questionnaires and therefore not covered by the *bioagel* data set.

Table 3: Number of children in the family in the specific age range

Bioage File	Number of Children in the Family						
	1 Child	2 Children	3 Children	4 Children	5 Children	6 Children	7 Children
Bioage01	3.066	885	121	19	3	2	1
Bioage02 (FiD only)	1671	232	18				
Bioage03	3434	930	136	19	4	1	
Bioage06	3093	990	115	10	2		
Bioage08a	2748	828	133	8			
Bioage08b	1923	562	87	5			
Bioage10	2692	567	89	4			
Bioage10b (FiD only)	785	169	7				
Bioage12	1440	159	5	1			
Bioage14	507	12					

The data also includes twins. *Bioage01* contains 103 pairs, *bioage02* 57 pairs, *bioage03* 109 pairs, *bioage06* 80 pairs, *bioage08a* 85 pairs, *bioage08b* 53 pairs, *bioage10* 72 pairs, *bioage10b* 24 pairs, *bioage12* 29 pairs of twins, and *bioage14* 8 pairs.

3   Topics and Variables

The *bioage* data contain information regarding

- Pregnancy and childbirth
- Child health
- Childcare situation
- Changes in living circumstances since birth of child
- Child's abilities
- Parenting experiences
- Expectations about the child's success in school
- Educational goals and aspirations of the parents
- Educational behavior of the parents
- Self-conception of the role of parents
- Self-report: independence and social experience
- Self-report: importance of social contacts
- Self-report of savings
- Self-report: interest in politics

The rules used to generate the variables from the questionnaires are consistent over the various *bioage* questionnaires. In the new integrated *bioage long* data set, the data is presented in "long" format, i.e. this dataset contains information from *bioage01*, *bioage03*, *bioage06*, *bioage08a* and *bioage08b, bioage10, bioage12* as well as *bioage14*. In addition, all variables from the FiD-Study are included as well: if there was a corresponding SOEP Variable, the data was combined into the SOEP variable; if there was no corresponding SOEP Variable or if the data was coded differently, the information is provided in separate variables with the suffix *_fid*.

A few variables are generated by combining the information from two or more variables. The present documentation provides detailed information on variables generated using information from other files. An overview of the specific variables and the rules by which they were generated is given in section 10.4 "Generated Variables".

4   Generated Variables

This section provides additional information on variables that have been generated from combinations of variables from other datasets than the parent questionnaires.

AGE

Variable label	**"Child's age in months"**
Variable format	2-digit integer
Comment:	This variable provides the child's age in months as a combination of month of birth and interview month. As the exact day of birth remains unknown, information is only an approximation and may vary by one month.
	Note that the information concerning year and month of birth from the "Mother & Child" and "Parent" questionnaires proved to be partially inconsistent. Therefore, as of the beginning of 2012 the child's age in months is computed using birth information from *ppfad*. For further information, refer to the documentation on *ppfad*.

PREGY

Variable label	**"Mother: pregnant at interview in survey year X"**
Variable format	4-digit integer
Bioage File	01
Comment	This variable is based on information from the previous year's individual questionnaire provided by the mother on her pregnancy status at the time of the interview. If pregnancy was reported (or was unknown) and a child was born, the year in which the interview took place is contained in BCPREGY. Hence this information is available only for those women in the sample for at least two years with a completed individual interview in the first year and a completed "Mother & Child" questionnaire in the second year. Please note that some mothers are not aware that they are pregnant in the early stages of pregnancy. The time of observation starts in survey year 2003 with the 2002 birth cohort.

PREGMO

Variable label	**"Mother: pregnancy month at interview"**
Variable format	2-digit integer
Bioage File	01
Comment	This variable is based on the exact month of birth (BCPREGMO), the duration of childbearing in weeks (BCSSW) and the interview month of the previous year's personal interview. Hence this information is available only for those women in the sample for at least two years. As the exact day of birth is unknown, this variable remains a close approximation.

PREBEG

Variable label	**"Spell Begin Pregnancy (Month, 01.83=1)"**
Variable format	3-digit integer
Bioage File	01
Comment	The variable BCPREBEG contains information on the beginning of pregnancy (i.e., the month of conception). Information is given in the regular SOEP spell format: values start with 1 for January 1983 (e.g., the earliest spell in *bioage01*, survey year 2010, is 304, which equals April 2008).
	The variable is based on the exact month of birth (BCPREGMO) and the duration of pregnancy in weeks (BCSSW). Accordingly, information is available only for women who completed the "Mother & Child Questionnaire" and for whom the duration of the pregnancy is known. Note that the month of conception may vary by one month as the exact date of birth remains unknown.

PREEND

Variable label	**"Spell End Pregnancy, Birth (Month, 01.83=1)"**
Variable format	3-digit integer
Bioage File	01
Comment	The variable BCPREEND contains information on the end of pregnancy (i.e. the month of birth). Information is given in the regular SOEP spell format: values start with 1 for January 1983 (e.g., earliest spell in *bioage01*, survey year 2010, is 304, which equals April 2008).
	This variable is based on the exact month of birth (BCPREGMO) and the duration of pregnancy in weeks (BCSSW).

SEX

Variable label	**"Gender of child"**
Variable format	1-digit integer
Comment	The sex of the child is not asked for in the "Mother & Child" and "Parent" questionnaires. Information on this variable stems from the *ppfad*.

SEXRESP

Variable label	**"Sex of respondent (parent)"**
Variable format	1-digit integer
Bioage File	08a, 08b
Comment	This variable tells whether the mother or father answered the respective questionnaire. The information for this variable comes from the *ppfad* file.