

SOEP Survey Papers

Series D – Variable Descriptions and Coding

SOEP-Core v33.1 – BIOSIB: Information on siblings in the SOEP

Josephine Kraft, Daniel D. Schnitzlein, and 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. Stefan Liebig, DIW Berlin and Universität Bielefeld

Dr. David Richter, DIW 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:

Josephine Kraft, Daniel D. Schnitzlein, and SOEP Group. 2018. SOEP-Core v33.1 – BIOSIB: Information on siblings in the SOEP. SOEP Survey Papers 582: Series D. Berlin: DIW/SOEP



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

© 2018 by SOEP

ISSN: 2193-5580 (online)

DIW Berlin

German Socio-Economic Panel (SOEP)

Mohrenstr. 58

10117 Berlin

Germany

soeppapers@diw.de

SOEP-Core v33.1 – BIOSIB: Information on siblings in the SOEP

Josephine Kraft, Daniel D. Schnitzlein, and SOEP Group

BIOSIB: Information on siblings in the SOEP

by Josephine Kraft and Daniel D. Schnitzlein

1 General description of the data set

BIOSIB provides information on siblings living within the SOEP households. The data set contains the person numbers of all siblings in an observed family. It includes information on their sex, their year of birth, the number of siblings, the individual's position within the birth order, and on the relationship between the observed siblings.

2 Sources of information on siblings in the SOEP

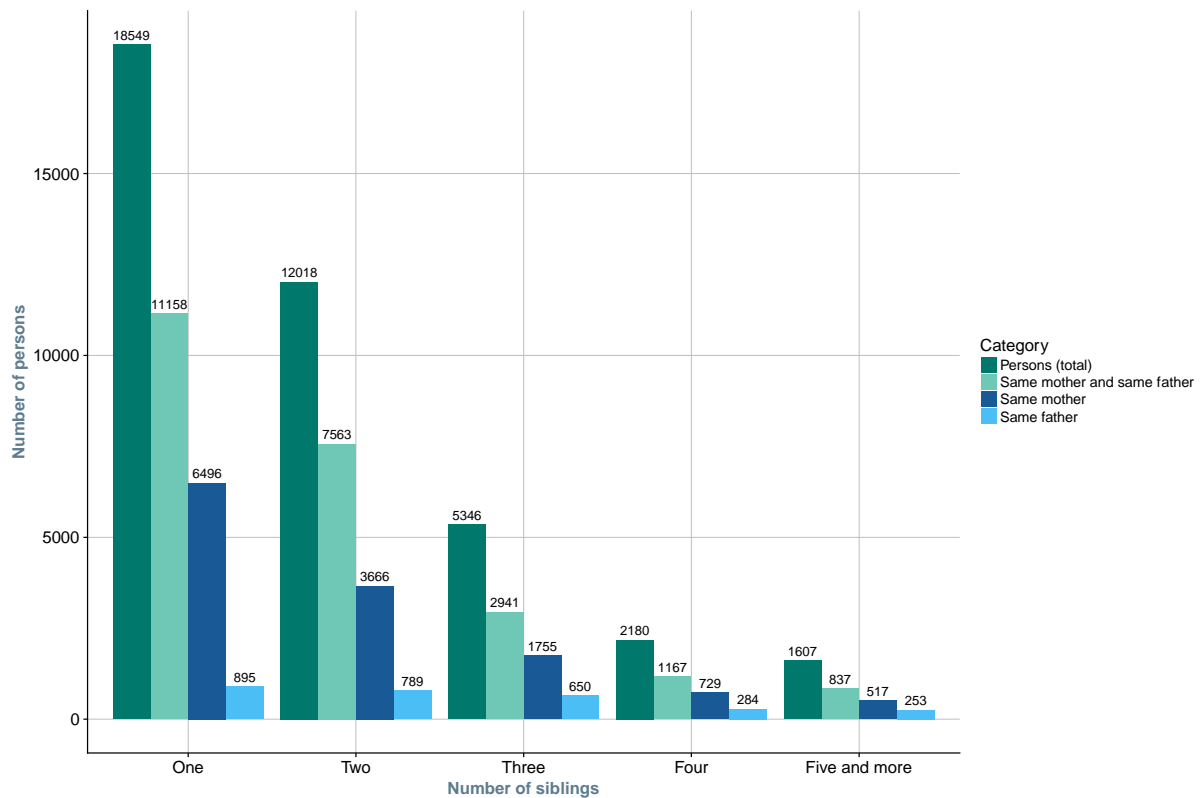
Information on siblings in the SOEP is available from three sources:

- First, the respondents are asked about their siblings in the biography questionnaire and the youth questionnaire. This information (for example *have or ever had siblings yes/no*, *number of sisters*, *number of brothers*) is stored in the file BIOPAREN (for detailed information please see the chapter on BIOPAREN).
- Second, in the years 1991, 1996, 2001, 2006, 2011 and 2013 all respondents were asked about their family relations. Among other questions on their family, the individuals were asked about their siblings (like above, *have or ever had siblings yes/no*, *number of sisters*, *number of brothers*) outside the SOEP household in 1991, 1996 and 2001. From 2006 on they were asked about all siblings within and outside the household. The information from these questions is stored in the \$\$p-files.
- Third, siblings within the SOEP households are observed directly. The aim of this file is to provide the never changing person ID of the siblings of each respondent as far as they can be identified in the SOEP. With this information it is possible to use the whole range of personal or household information for each sibling to carry out detailed sibling analyses. This file therefore adds to the information in BIOTWIN (for details see chapter on BIOTWIN), which provides the person IDs for all twins in the SOEP.

3 Overview on the number of siblings in BIOSIB

Figure 1 gives an overview on the information stored in BIOSIB. The data set contains 15,430 families with 39,735 individuals that have at least one sibling identified in the SOEP. 18,549 of them have one sibling, 12,018 have two siblings, 5,346 have three siblings, 2,180 have four siblings and 1,607 have five or more siblings identified in the data. As is apparent from Figure 1 in all cases most of the siblings are identified by having the same mother and the same father (for more details about the identification of siblings and the identification of biological siblings please see the description to SIBDEF1-SIBDEF11 on the subsequent pages).

Figure 1: Number of persons with information on their siblings, by number of siblings



Source: own calculation, based on SOEPv33.1.

4 Organization of the data in BIOSIB

Each row in the dataset represents one individual for which at least one sibling could be identified. Therefore a family with three siblings appears three times in BIOSIB, one time for each child. The person IDs of the siblings are ordered by birth order starting with the oldest sibling

List of variables

HHNR	Original Household Number
PERSNR	Never Changing Person ID
SIBPNR1 – SIBPNR11	Person Number of 1 st – 11 th Sibling
SIBDEF1 – SIBDEF11	Sibling Relation to 1 st – 11 th Sibling
FAMCOUNT	Family Counter
POS_SIB	Position in the birth order
NUM_SIB	Number of observed siblings in the SOEP
SEX	Gender of Individual
GEBJAHR	Year of Birth of Individual
SEXSIB1 – SEXSIB11	Gender of 1 st – 11 th Sibling
GEBSIB1 – GEBSIB11	Year of Birth of 1 st – 11 th Sibling

The variables HHNR, PERSNR, SEX, GEBJAHR, SEXSIB1-SEXSIB11 and GEBSIB1-GEBSIB11 are generated from the information stored in PPFAD. The newly generated variables SIBPNR1-SIBPNR11, SIBDEF1-SIBDEF11, FAMCOUNT, POS_SIB, and NUM_SIB are described on the next pages.

Variable	SIBPNR1 – SIBPNR11
Label:	Person Number of 1 st -11 th Sibling
Values:	(-1) No answer (-2) Does not apply (-3) Answer improbable
Description:	<p>The variables provide the never changing person IDs for the siblings of the individual identified by PERSNR. The sibling relationship is generated from the parent information in BIOBIRTH and BIOPAREN (for detailed information on these files please see the relevant chapters above). Two persons are defined as siblings if they report both, the same mother and father, only the same mother, or only the same father. This information on the sibling relationship is stored in SIBDEF1-SIBDEF11.</p> <p>In the case of inconsistent information on parents in BIOBIRTH and BIOPAREN, BIOPAREN was assigned the lowest priority.</p> <p>Please note, that BIOPAREN uses a social definition of parenthood based on cohabitation. In contrast, BIOSIB contains both biological (BIOBIRTH/BIOBRTHM) and social siblings with a higher priority on biological relations.</p>

Variable **SIBDEF1 – SIBDEF11**

Label: Sibling Relation to 1st-11th Sibling

Values:

- (-1) No answer
- (-2) Does not apply
- (-3) Answer improbable
- (1) Same mother (B); same father (B)
- (2) Same mother (B); same father (nB)
- (3) Same mother (nB); same father (B)
- (4) Same mother (nB); same father (nB)
- (5) Same mother (B)
- (6) Same mother (nB)
- (7) Same father (B)
- (8) Same father (nB)

Description: The variables provide the information on the sibling relationship between the individual identified by PERSNR and the respective sibling. Two siblings can have either, the same mother and the same father, only the same mother, or only the same father. The indicator further provides information if the identified parent is a biological (B) (indicator from BIOBIRTH) or non-biological parent (nB) (indicator from BIOPAREN). So for example, variable value (1) indicates that the two individuals share the same biological mother and the same biological father.

Variable	FAMCOUNT
Label:	Family Counter
Values:	(-1) No answer (-2) Does not apply (-3) Answer improbable
Description:	<p>The variable contains a non-systematic counter of families occurring in BIOSIB. All siblings (biological and non-biological), who belong to one family, are assigned the same value of FAMCOUNT. The variable can be used, for example, in multilevel analyses to define the family level.</p> <p>Note: In the case a family splits up, children from the new partnerships of the parents are no siblings. Children of the early partnerships are siblings to all children in the new partnership.</p>

Variable **POS_SIB**

Label: Position in birth order

Values:

- (-1) No answer
- (-2) Does not apply
- (-3) Answer improbable

Description: The variable contains the individual's position in the birth order of the observed siblings (biological and non-biological).

Variable **NUM_SIB**

Label: Number of siblings observed in SOEP

Values: (-1) No answer
 (-2) Does not apply
 (-3) Answer improbable

Description: The variable contains the total number of identified siblings, including the respondent, in the SOEP (biological and non-biological).