

SOEP-IS Group

Research Report

SOEP-IS 2023 - HPATHL: Household-related meta-dataset (long)

SOEP Survey Papers, No. 1476

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: SOEP-IS Group (2025) : SOEP-IS 2023 - HPATHL: Household-related meta-dataset (long), SOEP Survey Papers, No. 1476, Deutsches Institut für Wirtschaftsforschung (DIW), German Socio-Economic Panel (SOEP), Berlin

This Version is available at:

<https://hdl.handle.net/10419/316293>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-sa/4.0/>

1476²⁰²⁵

SOEP Survey Papers
Series D – Variable Descriptions and Coding

SOEP-IS 2023 – HPATHL: Household-Related Meta-Dataset (Long)

SOEP-IS Group

Running since 1984, the German Socio-Economic Panel (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/soepsurveyspapers>

Editors:

Dr. Jan Goebel, DIW Berlin
Dr. Christian Hunkler, DIW Berlin
Prof. Dr. Philipp Lersch, DIW Berlin and Humboldt-Universität zu Berlin
Dr. Levent Neyse, DIW Berlin and Berlin Social Science Center (WZB)
Prof. Dr. Carsten Schröder, DIW Berlin and Freie Universität Berlin
Prof. Dr. Sabine Zinn, DIW Berlin and Humboldt-Universität zu Berlin

Please cite this paper as follows:

SOEP-IS Group, 2025. SOEP-IS 2023 – HPATHL: Household-Related Meta-Dataset (Long). SOEP Survey Papers 1476: Series D – Variable Descriptions and Coding. Berlin: DIW Berlin/SOEP



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.
© 2025 by SOEP

ISSN: 2193-5580 (online)

DIW Berlin
German Socio-Economic Panel (SOEP)
Mohrenstr. 58
10117 Berlin
Germany

soeppapers@diw.de

SOEP-IS 2023 – HPATHL: Household-Related Meta-Dataset (Long)

SOEP-IS Group

2025

The file hpathl is part of a collection, which is released with doi:[10.5684/soep.is.2023](https://doi.org/10.5684/soep.is.2023).

Contents

1	Primary Key, Foreign Keys and Sample Information	3
	cid – Original Household Number, Case ID	3
	hid – Current Wave HH Number	3
	syear – Survey Year	3
	sample1 – Subsample Identifier	3
2	Survey History	4
	hnetto – Aktueller Befragungsstatus	4
3	Basic Demographic Information	5
	sampreg – Aktuelle Stichprobenregion	5
	hpop – Aktuelle Populationszugehoerigkeit	5
4	Weighting	5
	hbleib – Inverse Staying Probability	5
	hhrf – PRELIMINARY: Weighting factor	6

1 Primary Key, Foreign Keys and Sample Information

cid – Original Household Number, Case ID

Case Id - Household Source Identifier (fixed). The fixed household source id point to the first SOEP ancestor households to which current households can be traced back to. This variable is relevant for sampling information and calculating the weights.

hid – Current Wave HH Number

The members and configuration of a household with a given ID (HID) can change between years. Depending on the respective research issue it may be necessary to check alterations using PPATHL, where persons are assigned to a household in each year.

syear – Survey Year

1998	373
1999	388
2000	405
2001	428
2002	432
2003	443
2004	451
2005	463
2006	477
2007	476
2008	471
2009	1960
2010	2223
2011	1840
2012	4610
2013	6972
2014	7720
2015	4578
2016	8680
2017	4877
2018	4467
2019	4054
2020	4687
2021	2558
2022	3635
2023	13166

Survey year (Reference Year for survey instruments)

sample1 – Subsample Identifier

1	A German HH Head, West, 1984	0
2	B1 Turkish HH Head, 1984	0

3	B2 Yugoslav HH Head, 1984	0
4	B3 Greek HH Head, 1984	0
5	B4 Italian HH Head, 1984	0
6	B5 Spanish HH Head, 1984	0
7	C German HH Head, East, 1990	0
8	D1 Immigrant W. Germany (1984-1994), 1994/95	0
9	D2 Immigrant W. Germany (1984-1994), 1994/95	0
10	E1 Refreshment (PAPI), 1998	4941
11	E2 Refreshment (CAPI), 1998	4628
12	F1 Refreshment, 2000	0
13	F2 Refreshment, Additional Foreigner Addresses, 2000	0
15	G2 High Income, 2002	0
16	H Refreshment, 2006	0
...	(22 rows omitted)	71265
60	L2/3 FiD Screening, Low Income, 2010/11	0
61	L2/3 FiD Screening, Single Parent, 2010/11	0
62	L2/3 FiD Screening, Families with Several Children, 2010/11	0
63	L1 FiD Birth Cohort 2007, 2010/11	0
64	L1 FiD Birth Cohort 2008, 2010/11	0
65	L1 FiD Birth Cohort 2009, 2010/11	0
66	L1 FiD Birth Cohort 2010, 2010/11	0
-1	No answer	0
-2	Does not apply	0
-3	Implausible value	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question this year not part of survey	0

2 Survey History

hnetto – Aktueller Befragungsstatus

1	Successful HH Interview	50839
2	HH In Brutto Only	29995
3	unsuccessful, initial gross sample	0
-1	No Answer	0
-2	Does not apply	0
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question this year not part of Survey program	0

Wave-specific survey status: for HNETTO=1 there is a HH interview in the respective year, i.e. an entry in the file HL. For HNETTO=2 there is only an entry in the file HBRUTTO, but

the HH could not be interviewed (the reason for which is given by the variable HERGS in the file HBRUTTO). HNETTO=3 indicates unrealized Households of new samples in HBRUTTO-Files.

3 Basic Demographic Information

sampreg – Aktuelle Stichprobenregion

1	West-Germany	64041
2	East-Germany	14159
-1	No Answer	0
-2	Does not apply	76
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question this year not part of Survey program	2558

The variable SAMPREG tells whether a HH was located in West or East Germany in the respective year. This variable approximates also former regional differences within Berlin according to zip-codes.

hpop – Aktuelle Populationszugehoerigkeit

1	Private HH, German HH-Head	46156
2	Private HH, Foreign HH-Head	2180
3	Institutionalized HH, German HH-Head	244
4	Institutionalized HH, Foreign HH-Head	26
-1	No Answer	0
-2	Does not apply	32228
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question this year not part of Survey program	0

POP was derived from WUM2 in HBRUTTO as well as PNAT_V* and STELL_V* (nationality and relationship to head of household in PBRUTTO). Missing values were imputed based on the person's history. Thus, the only admissible missing value is -2, meaning not applicable. This variable is therefore particularly important, as it enters into the determination of cross-sectional weights. The variable corresponds with POP in PPATH/L.

4 Weighting

hbleib – Inverse Staying Probability

inverse probability weights

hhf – PRELIMINARY: Weighting factor

standard household weight. The weighting for the 2023 survey year is based on the previous year's marginal distribution. The updated marginal distributions of the Federal Statistical Office of Germany for 2023 were not yet available to us at the time of the data release. We therefore ask you to apply the weighting for 2023 with caution. The newly drawn sample in 2023 had to be integrated into the weighting with the information on the marginal distributions of the previous year, which can lead to additional inaccuracies.