SCP Documentation
German Social Cohesion Panel

SCP 2021/22 W1 Codebook PPATH: Person-Related Meta-Dataset (English)

German Social Cohesion Panel

Established in 2021, the German Social Cohesion Panel (SCP) is a wide-ranging representative longitudinal study of private households in Germany, carried out in collaboration of the Research Institute Social Cohesion (RISC) and the German Socio-Economic Panel (SOEP).

The aim of the SCP Documentation is to thoroughly document the survey's data collection and data processing.

▸ **Authors:** Olaf Groh-Samberg, Julian B. Axenfeld, Jean-Yves Gerlitz, Carina Cornesse, Martin Kroh, Holger Lengfeld, Stefan Liebig, Lara Minkus, Jost Reinecke, David Richter, Nils Teichler, Richard Traunmüller, Sabine Zinn

▸ **Contributors:** Cosima Adams, Anton Bochert, Martin Gerike, Josefine Kuhrmeier, Anna-Tabea Müller, Eric Nissen, Rainer Siegers, Hans Walter Steinhauer, Knut Wenzig, Julia Witton (Project Members), infas (Data Collector)

▸ **Publisher:** RDC-RISC
SOCIUM, University of Bremen
P.O. Box 330 440
28334 Bremen
Germany

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

▸ **DOI:** 10.60532/scp.2021_22.w1.v1

▸ **Website:** www.fgz-risc-data.de

www.diw.de

SCP 2021/22 W1 Codebook PPATH: Person-Related Meta-Dataset (English)

Contents

1 General Information

The PPATH dataset is designed to support longitudinal analysis linking personal information from various waves of the SCP. Each person participating in at least one survey wave is also in PPATH. PPATH also contains two persons who began the recruitment survey but did not answer any items and are excluded from the P, PGEN and PHRF datasets.

In some places in the documentation and in the data, year numbers are used, for example, for the names of variables and of the questionnaire instrument. These year numbers are always based on the field start of the data collection of the corresponding survey wave.

Data in this dataset may be missing due to different reasons. Missings are coded "[-1] No answer" if the respondent fails to provide valid information that is required for the variable. Missings are coded "[-2] Does not apply" if the information is missing due to the study design.

2 Identifiers and Sample Information

pid – Never Changing Person ID

2110000301		1
2110000302		1
2110000901		1
2110000902		1
2110001001		1
2110001201		1
2110001601		1
2110002001		1
2110002101		1
2110003201		1
2110003701		1
2110003901		1
2110004401		1
2110004402		1
2110004403		1
…	(16999 rows omitted)	16999
2113796701		1
2113796702		1
2113796703		1
2113797101		1
2113797102		1
2113797201		1
2113797301		1
2113797601		1
2113797801		1
2113797901		1
2113798501		1
2113798701		1
2113799101		1
2113799102		1
2113800001		1

The central individual identifier across time is PID, which is fixed over time (and of course datasets).

cid – Original Household ID

21100003		2
21100009		2
21100010		1
21100012		1
21100016		1
21100020		1
21100021		1
21100032		1
21100037		1
21100039		1
21100044		3
21100045		1
21100049		1
21100050		1
21100058		1
...	(13025 rows omitted)	16989
21137960		1
21137961		1
21137963		2
21137964		2
21137967		3
21137971		2
21137972		1
21137973		1
21137976		1
21137978		1
21137979		1
21137985		1
21137987		1
21137991		2
21138000		1

This identifier groups individuals into their original households at the start of the panel. That means, a person's CID is time-constant and will always relate them back to the household they initially belonged to, even if they moved out since.

psample_scp – Sample

1	[1] A 2021 Initial Sample	17029

Unchanging sample information

hid21_1 – Current Household ID 2021 Part 1

21100003		2
21100009		2
21100010		1
21100012		1
21100016		1
21100020		1
21100021		1
21100032		1
21100037		1
21100039		1
21100044		3
21100045		1
21100049		1
21100050		1
21100058		1
...	(13025 rows omitted)	16989
21137960		1
21137961		1
21137963		2
21137964		2
21137967		3
21137971		2
21137972		1
21137973		1
21137976		1
21137978		1
21137979		1
21137985		1
21137987		1
21137991		2
21138000		1

This identifier groups all individuals into their respective households at the time of the most recent wave (i.e. a person's HID can change over time, for example if an adult child moves out of their parents' home and starts their own household).

hid21_2 – Current Household ID 2021 Part 2

21100003	2
21100009	2
21100010	1
21100012	1
21100016	1
21100020	1
21100021	1
21100032	1

21100037	1
21100039	1
21100044	3
21100045	1
21100049	1
21100050	1
21100058	1
... (13025 rows omitted)	16989
21137960	1
21137961	1
21137963	2
21137964	2
21137967	3
21137971	2
21137972	1
21137973	1
21137976	1
21137978	1
21137979	1
21137985	1
21137987	1
21137991	2
21138000	1

This identifier groups all individuals into their respective households at the time of the most recent wave (i.e. a person's HID can change over time, for example if an adult child moves out of their parents' home and starts their own household).

3 Survey History

eintritt – Year First Contacted

2021	16666
2022	363

austritt – Year of Last Contact

2021	5218
2022	11811

4 Demographics

sex_scp – Sex

? What gender are you? *(from: scp/2021_2/Q50:pgender[])*

1	[1] Male	8281
2	[2] Female	8518
3	[3] Diverse	40
-1	[-1] No answer	190
-2	[-2] Does not apply	0
-3	[-3] Implausible value	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-7	[-7] Only available in less restricted edition	0
-8	[-8] Question not part of the survey program this year	0
-9	[-9] Don't want to answer [only CAWI]	0
-10	[-10] Break-off [only CAWI]	0

Missing values are replaced with valid values from the following survey wave.

sex_scp_flag – Flag for sex_scp

0	[0] Not changed since first inform.	16904
1	[1] Changed since first inform.	125

This variable indicates whether a valid value in PGENDER has changed at least once in a subsequent survey wave compared to the first valid entry. A change from a valid value to a missing value or vice versa is not displayed.

gebjahr – Year of Birth

? When were you born? // Year *(from: scp/2021_1/Q6:pbirthyear[])*

? In which year were you born? // Year of birth: *(from: scp/2021_2/Q49:pbirthyear[])*

1921	[1921] 1921 and before	2
1922	[1922] 1922	1
1923	[1923] 1923	3
1924	[1924] 1924	1
1925	[1925] 1925	5
1926	[1926] 1926	10
1927	[1927] 1927	7
1928	[1928] 1928	12
1929	[1929] 1929	13
1930	[1930] 1930	20
1931	[1931] 1931	29
1932	[1932] 1932	26
1933	[1933] 1933	36
1934	[1934] 1934	59
1935	[1935] 1935	81
...	(65 rows omitted)	15850
2001	[2001] 2001	243
2002	[2002] 2002	231

2003	[2003] 2003	160
2004	[2004] 2004	2
2005	[2005] 2005	2
-1	[-1] No answer	236
-2	[-2] Does not apply	0
-3	[-3] Implausible value	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-7	[-7] Only available in less restricted edition	0
-8	[-8] Question not part of the survey program this year	0
-9	[-9] Don't want to answer [only CAWI]	0
-10	[-10] Break-off [only CAWI]	0

Birth years 1921 and earlier were combined into one category for data protection reasons. Missing values are replaced with valid values from the following survey wave.

gebjahr_flag – Flag for gebjahr

0	[0] Not changed since first inform.	16837
1	[1] Changed since first inform.	192

This variable indicates whether a valid value in PBIRTHYEAR has changed at least once in a subsequent survey wave compared to the first valid entry. A change from a valid value to a missing value or vice versa is not displayed.

germborn – Born in Germany

[?] In what country were you born? *(from: scp/2021_1/Q15a:porigin[])*
[?] Since when do you mainly live in Germany? Since the year *(from: scp/2021_1/Q15c:pgermstay[])*

1	[1] Born in Germany or immigrated before 1949	15289
2	[2] Immigrated after 1948	1481
-1	[-1] No answer	259
-2	[-2] Does not apply	0
-3	[-3] Implausible value	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-7	[-7] Only available in less restricted edition	0
-8	[-8] Question not part of the survey program this year	0
-9	[-9] Don't want to answer [only CAWI]	0
-10	[-10] Break-off [only CAWI]	0

This variable's information is based upon PORIGIN. Respondents who stated that they immigrated to Germany before 1949 (PGERMSTAY) are assigned the value '[1] Born in Germany or immigrated before 1949'. For individuals with the value '[-1] No answer', no valid information could be generated.

immiyear – Year of Immigration to Germany

? In what country were you born? *(from: scp/2021_1/Q15a:porigin[])*
? Since when do you mainly live in Germany? Since the year *(from: scp/2021_1/Q15c:pgermstay [])*

1921	[1921] 1921 and before	0
1922	[1922] 1922	0
1923	[1923] 1923	0
1924	[1924] 1924	0
1925	[1925] 1925	0
1926	[1926] 1926	0
1927	[1927] 1927	0
1928	[1928] 1928	0
1929	[1929] 1929	0
1930	[1930] 1930	0
1931	[1931] 1931	0
1932	[1932] 1932	0
1933	[1933] 1933	0
1934	[1934] 1934	0
1935	[1935] 1935	0
...	(82 rows omitted)	1316
2018	[2018] 2018	54
2019	[2019] 2019	58
2020	[2020] 2020	41
2021	[2021] 2021	12
2022	[2022] 2022	0
-1	[-1] No answer	259
-2	[-2] Does not apply	15289
-3	[-3] Implausible value	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-7	[-7] Only available in less restricted edition	0
-8	[-8] Question not part of the survey program this year	0
-9	[-9] Don't want to answer [only CAWI]	0
-10	[-10] Break-off [only CAWI]	0

This variable's information is based upon PGERMSTAY. Respondents who were '[1] Born in Germany or immigrated before 1949' (GERMBORN) are assigned the value '[-2] Does not apply'. For individuals with the value '[-1] No answer', no valid information could be generated.

corigin_scp – Country of Origin

? In what country were you born? *(from: scp/2021_1/Q15a:porigin[])*
? In which other country were you born? *(from: scp/2021_1/Q15b:pothcount[])*
? Since when do you mainly live in Germany? Since the year *(from: scp/2021_1/Q15c:pgermstay [])*
? What other citizenship do you have? *(from: scp/2021_1/Q16b:pnatoth[])*

4	[4] Afghanistan	22
8	[8] Albania	11
10	[10] Antarctic	0
12	[12] Algeria	8
16	[16] American Samoa	0
20	[20] Andorra	0
24	[24] Angola	0
28	[28] Antigua and Barbuda	0
31	[31] Azerbaijan	2
32	[32] Argentina	4
36	[36] Australia	4
40	[40] Austria	38
44	[44] Bahamas	0
48	[48] Bahrain	0
50	[50] Bangladesh	0
...	(230 rows omitted)	16641
876	[876] Wallis and Futuna	0
882	[882] Samoa	0
887	[887] Yemen	0
894	[894] Zambia	0
900	[900] Kosovo	13
-1	[-1] No answer	286
-2	[-2] Does not apply	0
-3	[-3] Implausible value	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-7	[-7] Only available in less restricted edition	0
-8	[-8] Question not part of the survey program this year	0
-9	[-9] Don't want to answer [only CAWI]	0
-10	[-10] Break-off [only CAWI]	0

This variable's information is based upon POTHCOUNT and PORIGIN. Respondents who were '[1] Born in Germany or immigrated before 1949' (GERMBORN) are assigned the value '[276] Germany'. If information about the country of birth is missing and GERMBORN is '[2] Immigrated after 1948' valid information about a respondent's citizenship (PNATOTH) is used to generate CORIGIN_SCP. For individuals with the value '[-1] No answer', no valid information could be generated.

migback_scp – Migration Background

? In what country were you born? *(from: scp/2021_1/Q15a:porigin[])*

? Which citizenship do you have? // Another *(from: scp/2021_1/Q16a:pnation02[])*

? Next, we would like to know more about the origins of your parents and grandparents. // Were both your parents born in Germany? *(from: scp/2021_1/Q17:pmigback01[])*

? Next, we would like to know more about the origins of your parents and grandparents. // Were all of your grandparents born in Germany? *(from: scp/2021_1/Q17:pmigback02[])*

1	[1] No Migration Background	12016

2	[2] Direct Migration Background	1780
3	[3] Indirect Migration Background	2292
4	[4] Migration Background, Type unclear	6
-1	[-1] No answer	935
-2	[-2] Does not apply	0
-3	[-3] Implausible value	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-7	[-7] Only available in less restricted edition	0
-8	[-8] Question not part of the survey program this year	0
-9	[-9] Don't want to answer [only CAWI]	0
-10	[-10] Break-off [only CAWI]	0

This variable provides information on migration background. It considers the country of birth and citizenship of the individual and his/her migration history. Direct/indirect migration background indicates whether the respondent immigrated by him/herself or has a non-German citizenship (direct migration background) or if she/he is of migrant origin but born in Germany (indirect migration background). Respondents assigned the code 4 also have a migration background, but we do not know whether they immigrated by themselves because the information on their country of birth is missing. Code '[1] No Migration Background' includes respondents of German nationality born in Germany whose parents and grandparents were also born in Germany. For individuals with the value '[-1] No answer', no valid information could be generated, mostly due to missing information on the parents' or grandparents' country of birth.

sampreg_scp21_1 – Current Sample Region 2021 Part 1

1	[1] Western Germany	11063
2	[2] Eastern Germany	5212
3	[3] Berlin	754

This variable indicates in which sample region households were located in the given survey wave: Western Germany, Eastern Germany, or Berlin.

sampreg_scp21_2 – Current Sample Region 2021 Part 2

1	[1] Western Germany	11063
2	[2] Eastern Germany	5212
3	[3] Berlin	754

This variable indicates in which sample region households were located in the given survey wave: Western Germany, Eastern Germany, or Berlin.