November 2023

Old Age in Germany (D80+):
User Manual of SUF D80+, Version 1.0

Stefan Stuth

TABLE OF CONTENTS

1. OLD AGE IN GERMANY (D80+) - A STUDY ON OLD AGE IN GERMANY

The study Old Age in Germany ("Hohes Alter in Deutschland") (D80+) was funded by the Federal Ministry for Family Affairs, Senior Citizens, Women and Youth (BMFSFJ) and conducted by the cologne centre for ethics, rights, economics, and social sciences of health (Ceres) in cooperation with the German Centre of Gerontology (DZA). The study focussed on very old people in Germany who are 80 years or older. The aim of the study is to systematically survey people in old and very old age. The study thus closes a gap in empirical research by using a sampling plan without limits on higher ages that includes communal accommodation (e.g., retirement homes).

In order to draw a representative picture of the very old population and their living conditions, a representative sample of the resident population aged 80 and over was surveyed. This sample included people in both private and nursing homes. The respondents were asked comprehensive questions about their objective and subjective living conditions. In addition, test procedures were used to assess cognitive performance. Proxy interviews were conducted in order to include very old people in the survey who were no longer able to participate in person for health reasons. The combined survey of very old people in private households and in care homes and the use of proxy interviews offer a unique representative data source for interdisciplinary research into the life situations of old and very old people in Germany.

The microdata from the Study on Old Age in Germany (D80+) are available free of charge for scientific research for non-commercial purposes. The Research Data Centre of the DZA provides researchers with access to the anonymised D80+ dataset and the documentation materials (https://www.dza.de/forschung/fdz).

The D80+ study builds on the NRW80+ study in terms of methodology and content. The study on old age in Germany was originally planned as a face-to-face interview (CAPI). Due to the coronavirus pandemic and the particular vulnerability of the elderly to COVID-19, the planned survey design had to be completely changed. The survey programme was divided into two modules: Module one records essential information that is easy to collect (Module 1). The second module collects further information that is more complex to query (Module 2). Module 1 was collected in writing (PAPI), Module 2 by telephone (CATI) around six months later. The survey of both modules took place from the end of 2020 to 2021. The Institut für angewandte Sozialwissenschaft GmbH (Infas), based in Bonn, was responsible for the fieldwork, data collection and data weighting.

1.1 Sample

The population of the D80+ sample is based on all registered persons in Germany born before 1 March 1940. The sample was drawn in two stages: 1) Drawing of a population-proportional community sample. 2) Drawing of a disproportionately stratified population sample from the community sample (stratified by age and gender with 6 stratification cells). For the first survey

phase, the written survey (Module 1), 40,209 people were contacted. 10,360 valid questionnaires were returned. The second survey phase was aimed at all persons in the sample of the written survey for whom a telephone number was available and who had not explicitly refused to participate. People who had already taken part in the written survey were asked the questionnaire of Module 2 by telephone. In a few cases, additional questions were asked to clarify ambiguities in the written questionnaire. People who did not take part in the written survey were asked the questions from Module 1 and Module 2 by CATI in the second phase. A total of 3,260 people took part in phase two. Of these 3,260 people, module 1 was also surveyed in CATI mode for 218 people (people who did not take part in the written survey in phase 1). Of these 218 people, 27 did not take part in Module 2, meaning that information on Module 2 is available for a total of 3,233 people.

The SUF D80+ dataset contains information on the study participants from the following phases and modules:

Sample D80+	Module 1	Module 2
Phase 1 (PAPI)	10.360	-
Phase 2 (CATI)	218	3.233
Total	10.578	3.233

1.2   Special features of the sample

The population sample in D80+ does not suffer from restrictions that are common in most other surveys:

1) D80+ uses a sample without an upper age limit. However, there is a lower age limit: all respondents must have been at least 80 years old at the time of the draw or have been born before 1 March 1940. This leads to the following distribution of respondents across three age groups: 80-84 year olds make up around 48 per cent of the D80+ sample, 85-89 year olds make up 31 per cent and respondents who were 90 years or older make up around 21 per cent of the D80+ sample.

2) D80+ uses a sample that is not limited to specific forms of housing. This means that institutional addresses of care homes for the elderly were included in the sampling process. Of all respondents in D80+, 83 per cent stated that they live in a private flat or house. 5.6 per cent of the sample live in an institutional form of accommodation. The remainder of respondents live in mixed forms of housing that are neither clearly private nor clearly institutionalised.

3) Surveys can generally only include people who are healthy enough to be able to take part in the survey. Particularly in old age, the proportion of people who are no longer able to provide

their own information due to illness or disability increases. Excluding them would mean accepting a positive bias, at least for health and mental well-being characteristics. In order to avoid this potentially positive bias, proxy interviews were allowed. (For more information, see chapter 3.)

1.3 Accommodation and care

What is special about the D80+ study is that, for the first time, people who were accommodated in residential care were explicitly included in the sample. The study records the type of accommodation and differentiates between institutional and private accommodation. Private accommodation is not to be equated with private rented accommodation. Private accommodation includes private flats, multi-generation homes and outpatient assisted living. Residential homes for the elderly or residential complexes with assisted living flats and senior residences are considered private as long as the residents do not receive full inpatient care. Retirement homes and care facilities are considered institutional accommodation. Residential homes for the elderly or residential complexes with assisted living flats and senior citizens' residences are categorised as institutional accommodation if the residents receive full inpatient care. A total of 587 people were categorised as institutional accommodation.

Full inpatient care refers to the fact that extensive care services are required on a continuous basis. The exact definition given to the interviewees is: full inpatient care is defined as living in a nursing home/care home for the elderly**,** requiring care **and** having access to care, support and nursing by a carer 24 hours a day. Private care, outpatient care or care provided by a care service are **not** considered to be fully inpatient care.

In addition to the original queries in the variables, a generated variable is also available (wohnf1_g). This variable attempts to resolve inconsistencies in response behaviour. According to this generated variable, 1,018 respondents were in full inpatient care. The fact that the resolution of the inconsistencies was only partially successful is shown by the intersection with the type of accommodation: 456 people were categorised as receiving full inpatient care **and** as living in a private type of accommodation. The source of the inconsistencies cannot be conclusively clarified, as the survey mode (self-completion questionnaire) does not allow any follow-up questions. It is conceivable that full inpatient care is sometimes also provided in a private setting, that the term full inpatient care was understood differently by the respondents than by the research team or that the questionnaire was not completed with the necessary care by everyone.

1.4 Subject areas

Both modules collected information on the following topics:

– Socio-demographics
– Living
– Family and household
– Financial situation
– Dealing with age
– Health
– Everyday organisation and lifestyle
– Technology utilisation
– Value system
– Biography
– Social integration (Module 1 only)
– Well-being and life satisfaction (Module 1 only)
– Personality (Module 2 only)
– Critical life events (module 2 only)
– Impairment due to Covid19 pandemic and measures (module 2 only)
– Dementia (Module 2 only)

Some subject areas are covered in more detail in Module 1 (e.g. socio-demographics and family and household). Other subject areas are surveyed in more detail in Module 2 (e.g. use of technology or financial situation). Some subject areas are only recorded in one of the two modules. These include, for example, social integration or well-being and life satisfaction, which were only recorded in Module 1. Questions on personality and critical life events, on the other hand, were only asked in Module 2. In Module 2, a complex survey instrument was also used in the health area to measure dementia.

2 RESPONSE RATE AND CHECKS

In phase 1, the written survey, 10,677 interviews were realised from a gross sample of 40,209 addresses. This corresponds to a response rate of almost 27 per cent. Phase 2 (telephone survey) required the availability of a telephone number. Furthermore, people who had died in the meantime or refused (additional) interviews were excluded from the gross sample of phase 2. These conditions were met by 6,265 of the 10,677 people who completed the questionnaire in phase 1. Valid interviews could be realised for 37.9 percent of the 6,265 people (N = 2,377). To increase the response rate among the respondents without a known telephone number, they were contacted and asked to participate in the telephone interview with an enclosed cash incentive of 5 euros and the promise of a further 20 euros. In this way, 703 telephone interviews could be realised from additional 4,289 participants in the written survey. Of the 29,532 people

who did not take part in phase 1, 8,568 people fulfilled the conditions for participation (telephone number, not deceased, no refusal). However, interviews could only be realised for 218 of these cases.

Tests for quality and validity reduced the number of cases: Interviews in which the respondents answered less than 50% of the contents of Module 1 were excluded, as were cases in which it was suspected that the interviews were completed by a person other than the target person. This means that the SUF contains 10,578 cases that completed Module 1 and 3,233 cases that completed Module 1 and Module 2.

3   PROXY INTERVIEWS (PROXIES)

While no proxy interviews were permitted for the written part of the survey (phase 1), it was possible to nominate a proxy for the telephone survey (phase 2). If feedback from phase 1 indicated that health reasons prevented participation in the survey, it was asked whether relatives of the target person could be contacted. This person should be able to provide information on behalf of the target person. Relatives, friends, legal representatives and caregivers of the target person were considered for this purpose. If a representative/proxy agreed to conduct the interview, the target person also had to give their consent. Consent could be given over the phone or in writing using a consent form that was mailed to the target person before the interview took place.

Data from proxy interviews is available for a total of 363 people: For 2 persons who participated only in Module 1 but not in Module 2), for 322 respondents who used proxy interviews in Module 2 (participation in Module 1 by questionnaire - no proxy interviews were permitted here) and for 39 respondents who used proxy interviews in Modules 1 and 2 (participation in Module 1 by telephone interview).

4   DATA WEIGHTING OF THE MAIN SURVEY 2020/21

The weighting factors were calculated separately for module one and module two. A design weight was created for each module. This records the inverse of the product of the overall selection probability of a person with the probability that a person will be included in the gross sample for the field assignment. The design weight was adjusted for non-response in a second step. The default models used for this included the following variables:

Non-response model module 1	non-response model module 2
Federal state	CATI1
Municipality size class	Municipality size class

Residential accommodation	Residential care
Telephone number available	Household size
Age x gender	Degree of care
	Education x Gender
	General health assessment
	Age group of the design group

With the help of the non-response models selective participation probabilities were calculated and multiplicatively linked to the design weight as an inverse. The weights created in this way were then calibrated or post-stratified in a third step. With the help of IPF (Iterative Proportional Fitting), the distribution of the sample was adjusted to the known distribution of the population that was based on information from the 2019 microcensus. The following information from the microcensus was taken into account Federal state, municipality size class, age group, gender, household size, marital status and home accommodation.

A total of 4 weighting factors are available: The variables "weight" and "desgew" refer to information that was collected in module 1. The variables cati2desgewnorm and catii2kal refer to information collected in module two. By default, only the complete weights (weight and catii2kal) are delivered in the SUF. Anyone interested in the design weights can request them separately from the FDZ.

5   DATA FORMATS

The Scientific Use File (SUF) of the D80+ study is available both in SPSS format and in Stata data format. If there are any version-related problems with importing the data, please contact the FDZ-DZA directly (fdz@dza.de). All data and the corresponding documentation are available in German and English.

The Stata dataset is delivered with two different label sets (German and English). To switch back and forth between the two languages for variable and value labels, use the Stata command `mlanguage set en` or `mlanguage set de`. For SPSS, the German and English labels are stored in two separate data sets.

6   CLASSIFICATION OF VARIABLE NAMES AND MISSING CODES

Variables originating from module 2 have the suffix _c2. Open-ended responses were recoded by the D80+ team. Variables with these recoded open-ended responses have an x in the variable name. The naming of the variables followed their content. They therefore have descriptive names that generally provide a rough insight into their content.

The assignment of missing codes is based on the DEAS.

The following table describes the codes for missing values that occur in the data set.

Value SPSS	Value Stata	Label
-1	.a	Refused
-2	.b	Don't know
-3	.c	Does not apply (overfiltered)
-4	.d	Not surveyed (sample) *[questions in module 2 if respondents only answered module 1]*
-6	.f	Information missing
-7	.g	Deleted in Data preparation
-8	.h	Multiple answers
-9	.i	Not possible (visual difficulties)
-10	.j	Not possible (problems with writing)
-11	.k	Missing original values [quality assurance - there were not enough valid data in the original variables to form the scale or the aggregate]
-12	.l	Not possible (hand gripping force)

The missing codes were formatted in the data set as user-defined system missings.

7   MODIFICATION OF DATA IN THE SUF TO MAINTAIN DE FACTO ANONYMITY

In order to preserve the anonymity of the study participants, individual variables were not included in the Scientific Use File (SUF) D80+ and the values of a number of variables were coarsened.

- Among other things, all variables relating to the spatial context were deleted, which might have made it possible to determine the respondent's municipality of residence. The information on the respondent-'s birthday and -month was also removed.

- Variables with open-ended information were deleted. The open-ended information was recoded by the D80+ study team. The recoded information is available to the data users.

- Rare family demographic characteristics have been coarsened (e.g. country of origin of one parent).

- For a number of variables, values with a small number of cases were summarised into categories (so-called "topcoding"; e.g. for household size, number of children/grandchildren, years). These are labelled to indicate which values were combined.

8   THE COMPLEX SURVEY DESIGN OF THE D80+

The D80+ was originally planned as a face-to-face survey. For this reason, the sample is not only stratified, but also lumped. In terms of logistics and costs, it is almost impossible to cover the whole of Germany. The travel and organisation costs for conducting the personal interviews would be disproportionately high. A size-proportional sample of municipalities from all federal states was drawn. Size-proportional means that larger municipalities had a higher probability of being drawn than smaller municipalities. As a result, 461 municipalities were drawn.

The D80+ is disproportionately stratified at the second selection stage (after the persons have been drawn from the population registers of the drawn municipalities). Six sample cells were formed based on the criteria of age and gender.

If this complex sampling design is not taken into account (in particular the clustering), the respondents are categorised as independent observation units. However, respondents from the same communities may be similar to each other (e.g. respondents from poor/rich communities are likely to be very similar in their income). If clustering is not taken into account, the variance is underestimated and statistically significant effects/relationships might be found that would not have been found if the design had been correctly taken into account.

For this reason, additional variables are provided in the D80+ in order to be able to communicate the complexity of the survey design to the statistical analysis programmes. In Stata, this can be done with the svy command-suite. The following characteristics are provided in the data set for this purpose: PSU_rand (Primary Sampling Unit), fallnr (Secondary Sampling Unit), and strata (sample cells used for the stratification of the SSUs).

The svyset command in Stata, which should be used to define the survey design features of the D80+, looks like this:

svyset PSU_rand, vce(linearised) singleunit(scaled) || fallnr [pw=weight], strata(strata)

The appropriate weight must be selected or created independently in accordance with the analysis design. If the analyses only contain information from Module 1, the weight "weight" should be selected. As soon as information from Module 2 is used, the weight "catii2cal" should be used and specified in svy-Set. All subsequent descriptive analyses should then include the suffix -svy-, as the complex survey design of the D80+, which was defined using the svyset command, is then taken into account. For more complex analyses such as regressions, the svy command does not necessarily have to be used. Instead, the variable PSU_rand should be used to calculate clustered standard errors. Furthermore, the values of the variable strata should be included in the analysis as dummies and the appropriate weighting factor should be taken into account too. In SPSS, the complex sample design can be included using a plan file that has PSU_rand as the cluster variable and either gewicht (only information from Module 1) or catii2kal ((also) information from Module 2) as the sample weighting variable.

9   DOCUMENTATION

Various documentation materials can be downloaded from the documentation section of the FDZ-DZA website. The materials include the complete survey instrument, the infas method report, the code book and the detailed scale documentation. These documents are also available in English. However, the English documentation was only produced automatically (by DeepL) and without additional checking or revision.

All known publications based on the D80+ are also listed on the FDZ-DZA website. The list is also available as a PDF document. If you would like to report your own publications based on the D80+ for this documentation, please contact fdz@dza.de.

10   REGIONAL DATA

The SUF of the D80+ cannot be analysed at regional level.