User Guide Release 7-1-0

WHAT'S NEW?!

We are happy to offer you a few new features in the SHARE-RV Release 7-1-0!

- One more reporting year was added to the administrative data: the VSKT is available until 2017 and the RTBN until 2018.
- Altogether administrative data are available for 4383 SHARE respondents – 4330 in the VSKT and 3060 in the RTBN.

CONTENT

Stand: 15.04.2020

1.   ABOUT SHARE-RV

1.1   WHAT IS SHARE-RV ABOUT?

SHARE-RV is a cooperation project between the Munich Center for the Economics of Aging (MEA), the Max Planck Institute for Social Law and Social Policy, and the Research Data Center of the German Pension Insurance ([FDZ-RV](#)[1]) in Berlin. It stands for the direct linkage of survey data of the *Survey of Health, Ageing and Retirement in Europe* ([SHARE](#)[2]) with administrative records of the "Deutsche Rentenversicherung", German Pension Insurance ([DRV](#)[3]). More precisely, it is the linkage of the German subsample of SHARE survey data and selected administrative records provided by the FDZ-RV. The project started as a pilot study in 2009 as part of the 3rd SHARE wave and was primarily funded by the Volkswagenstiftung. Thanks to the follow-up funding of the Forschungsnetzwerk Alterssicherung (FNA), the project continues at least until SHARE`s nineth wave of data collection.

The goals of the project SHARE-RV are to link survey data from SHARE with data from the DRV and to make them available to the entire research community. Hence, SHARE-RV allows the investigation of connections between various aspects of respondents´ lives and their working history or their socio-economic status in later life.

1.2   HOW IS SHARE-RV ORGANIZED?

The project is based on a direct linkage, meaning that the data of exactly the same person (here German SHARE respondents) were linked using the respondents' Social Security Number (SSN) as a unique identifier. When linking survey data with the records of the DRV, respondents' written consent is required. Respondents are asked for consent during the regular SHARE interview by the interviewer, which is documented on a separate consent form. The consent form collects respondents' SSN as well as some basic demographics which are needed to check the SSN. Filled-in consent forms without the respondents' signature cannot be considered for the linkage. If respondents are not able or not willing to provide their SSN, it is possible to find it out based on basic

---

[1] www.fdz-rv.de
[2] www.share-eric.eu
[3] www.deutsche-rentenversicherung.de

Stand: 15.04.2020

demographics provided by the respondents. In order to be absolutely sure that the correct records are linked, a lot of effort is put in checking and correcting the SSN.

1.3 WHO IS ELIGIBLE FOR RECORD LINKAGE?

In SHARE Wave 3, all German respondents were asked to consent to record linkage. In wave 5 and wave 6, all respondents who have not consented before or were new in the SHARE sample were asked to consent. The corresponding administrative data of the consenters are drawn yearly by the DRV at fixed dates.

Please note that the survey and the admin data are not necessarily representing and accordingly reporting the same year (see Chapter 4.3 reporting year for more details).

1.4 FOR WHICH SHARE RESPONDENTS ARE ADMINISTRATIVE DATA AVAILABLE?

Firstly, the DRV is only allowed to provide the records if respondents consented explicitly during the interview. Secondly, all inhabitants with at least one spell of activity that is subject to mandatory pension insurance are included in the administrative records of the DRV. So, what does that mean concretely and which activities are subject to mandatory pension insurance?

1. Who is compulsorily insured?

Generally, all employees who work in Germany are subject to mandatory pension contributions in the German Pension Insurance. Nevertheless, there are some exceptions to this rule, some special arrangements for certain occupational groups, and some activities/situations that go along with compulsory insurance independently of the employment status.

Additional to employees, also some other groups are subject to mandatory pension insurance, these are[4]:

- apprentices/trainees
- persons who raise children
- non-commercial caregivers
- persons with disabilities

---

[4] https://www.deutsche-rentenversicherung.de/DRV/DE/Rente/Arbeitnehmer-und-Selbststaendige/02_Pflicht-und-freiwillig-Versicherte/pflicht-und-freiwillig-versicherte_Inhalt_01_selbstaendig_und_pflichtversichert.html

- persons in military or voluntary service
- persons who receive compensatory service (e.g. unemployment or sickness benefit)
- Students with marginal jobs (including exceptions)

Although the rule is that self-employed persons do not have to be insured in the DRV, there are, furthermore, some self-employed persons who are compulsorily insured[5]:
- craftsmen
- teacher, midwives, educator, caregivers
- artists and publicists
- self-employed with only one customer
- pilots and coast boatmen

→ SHARE respondents who fall into these groups are hence included in the administrative RV records that are available for linkage, provided they have given consent for that.

For people in marginal employment, there are special regulations[6] that depend on the duration of the employment, the monthly salary as well as on the person's choice. For the so-called "Mini-Jobs" (max. 450€/month) there are two regulations: Short-term Mini-Jobs, that is up to three month, are not compulsorily insured, but marginal employments for more than three months have to be insured. The so-called "Midi-Jobs" (450,01€-850€/month) are subject to mandatory social insurance, but employers do not have to pay the full social insurance contributions. In both cases – Mini-Job and Midi-Job – it is possible to ask for an exemption from compulsory pension insurance. For detailed information, please see: "Geringfügige Beschäftigung und Beschäftigung in der Gleitzone"[7]. In the data of the German Pension Insurance marginal employment is

---

[5] https://www.deutsche-rentenversicherung.de/DRV/DE/Rente/Arbeitnehmer-und-Selbststaendige/03_Selbststaendige/selbststaendige_node.html#doc58496146-f190-4116-8e77-d7626a594659bodyText1

[6] https://www.deutsche-rentenversicherung.de/DRV/DE/Rente/Arbeitnehmer-und-Selbststaendige/04_Mini-Midi-und-Nebenjobs/mini-midi-und-nebenjobs_node.html

[7] https://www.bmas.de/SharedDocs/Downloads/DE/PDF-Publikationen/a630-geringfuegige-beschaeftigung.pdf?__blob=publicationFile

Stand: 15.04.2020

documented no matter if it is compulsorily insured or not. The difference can only be seen in the earning points.

2. Who isn't obligated to be insured?

There is no statutory duty to be insured in the DRV for[8]

- civil servants
- judges
- professional militaries
- statutory members of spiritual cooperatives
- marginal employed persons
- old-age pensioners after the month, in which the standard retirement age was reached
- self-employed and freelancers, for whom there is no duty for insurance

➔ SHARE respondents who fall into these groups are normally not included in the administrative records of the DRV that are available for linkage, except they had at least one month of insured activity.

Members of some occupational groups, actually, can be exempted from compulsory insurance as well. For members of professional associations ("Berufsständische Kammern") and liberal professions, special pension schemes were founded. These professional pension schemes are completely comparable to the DRV and also part of the 1st pillar of the pension system in Germany. The professional pension schemes for the liberal professions only exist for the members of specific professional associations ("Berufskammern") as there are[9]

- physicians, veterinaries, dentists
- pharmacists
- architects
- notaries and lawyers
- tax consultants and chartered accountants
- psychological psychotherapists

---

[8] https://www.deutsche-rentenversicherung.de/DRV/DE/Rente/Arbeitnehmer-und-Selbststaendige/02_Pflicht-und-freiwillig-Versicherte/pflicht-und-freiwillig-versicherte_Inhalt_01_selbstaendig_und_pflichtversichert.html

[9] http://www.abv.de/english-summary.html

Stand: 15.04.2020

The insurance in the professional pension schemes, however, does not apply automatically. Firstly, every single person has to ask for the exemption from compulsory insurance at the DRV. Secondly, since 2012 only persons who actually work profession-specific can be insured in a pension scheme. Thirdly, the exemption from compulsory insurance has to be checked again with every change in job[10]. However, the exemption also lasts for periods of e.g., unemployment, child-raising etc.

➔ SHARE respondents who fall into these groups are normally not included in the administrative records of the DRV that are available for linkage, except they did not work profession-specific and could not be exempted.

Anyway, there are persons who have an account at the German Pension Insurance but just with a very few entries. This can happen when there was a period that was subject to mandatory pension insurance and the status changed later to e.g. civil servant. In the German Pension Insurance already one month of compulsory insurance counts. Unfortunately, we cannot tell from the data if this is the case or not. But in some cases, the SHARE data can help: From the variable "job type" for example, you can find out if a respondent started to work in an occupation that is not subject to compulsory insurance. The following table shows the relevant occupations of the respondents for whom - despite their consent - no admin data is available.

Table 1 Occupation relevant to compulsorily insurance

Occupation	Frequencies	Percent	Cum.
Employee	37	25.34	
Civil Servant	76	52.05	77.40
Self-employed	23	15.75	93.15
not clear	10	6.85	100.00
Total	146	100.00	

Source: FDZ-RV – SUF SHARE-RV 7-1-0.

---

[10] Dr. Katharina Roth: Berufsständische Versorgung und Befreiungstatbestände in der gesetzlichen Rentenversicherung (https://www.deutsche-rentenversicherung.de/Schwaben/DE/Presse-und-Fachinformationen/Experten/Bayerische-Fachinformationen/2014_archiv/02-2014_Berufsstaend_Versorgung.pdf?__blob=publicationFile&v=1).

Some respondents, however, fulfill the conditions for linkage and record availability (written consent and spells of compulsory insurance), but their administrative data are not available anyway. This is the case if, for example, short-time insured persons ask for reimbursement of their contributions. Another reason for missing records could be that entitlements were accumulated in the former German Democratic Republic and the respondents retired before their records could be transferred to the new system.

The availability of administrative data that could be linked to the corresponding records is summarized in wave-specific linkage rates (see Table 2). This rate is calculated as the number of respondents in a given wave whose data can be linked, divided by all respondents who participated in that wave. The linkage rates considered in Table 2 refer to the most recent release version only. As the number of respondents differs from wave to wave due to attrition and new persons being interviewed, the linkage rate differs from wave to wave, too.
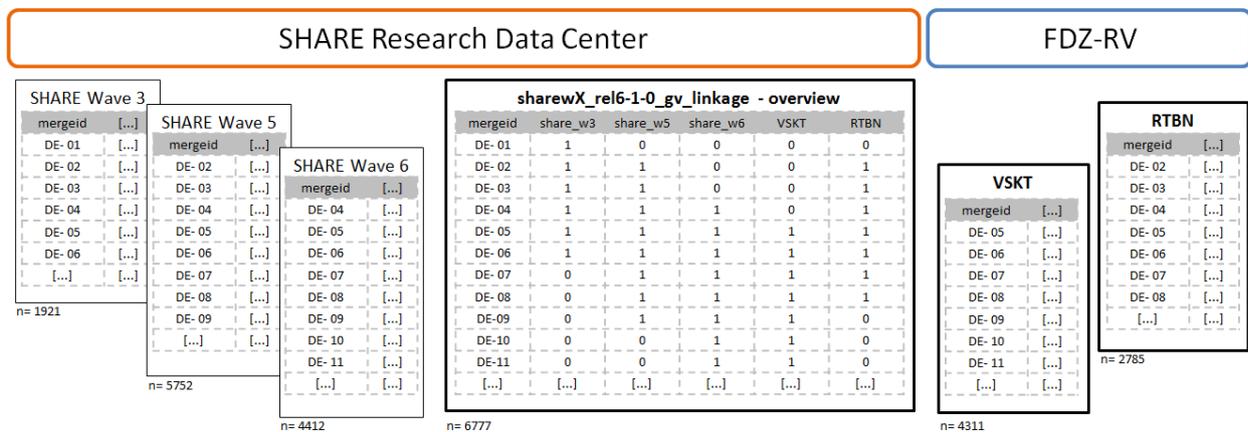
Table 2 Linkage rate by wave

	Wave 3 (Rel. 7-1-0)	Wave 4 (Rel. 7-1-0)	Wave 5 (Rel. 7-1-0)	Wave 6 (Rel. 7-1-0)	Wave 7 (Rel. 7-1-0)
No. of interviews in SHARE	1919	1619	5751	4412	3821
No. of linkable cases	1171	1105	4005	3352	2900
Record linkage rate in %	61,0	68,3	70	76	75,9

2. DATA

All data sets (SHARE and the administrative records) will be updated regularly, resulting in new release versions. A new release version of the administrative data will be published when either a new release of the SHARE data or a new version of the administrative records becomes available.

Hence, users have to ensure that they use the most recent release of the datasets[11]. The SHARE survey data and the administrative data sets are provided separately by the corresponding research data center. Figure 1 gives a detailed overview of the two dataset providers and the available data sets.

Figure 1 Datasets

2.1 SHARE

2.1.1 SURVEY DATA

SHARE is a multidisciplinary and cross-national panel database. It provides micro data on health, socio-economic status as well as social and family networks of more than 120000 individuals aged 50 or over and their (younger) partners. To date, SHARE has collected more than 297000 interviews in six panel waves on current living circumstances and one wave on retrospective life histories (SHARELIFE). The first wave was collected in 2004/2005, the second in 2006/2007, SHARELIFE in 2008/2009, the fourth wave mainly in 2011, the fifth wave in 2013 and the sixth wave in 2015. SHARE's seventh wave took place in 2017 and so far, 29 countries have participated in SHARE. The Wave 7 questionnaire contains a SHARELIFE module for all respondents who did not participate

---

[11] For an overview of SHARE releases see: http://www.share-eric.eu/data-documentation/share-data-releases.html

in Wave 3 (first SHARELIFE wave), as well as a standard module for all respondents who already answered a SHARELIFE interview. SHARELIFE focuses on people's life histories including all important areas of the respondents' lives, ranging from partners and children to housing and work history to detailed questions on health and health care. However, not all countries participated in every wave and the timing of data collection can also slightly differ between countries[12].

The German SHARE data collection started with the first wave. Until Wave 7 approximately 8800 individuals have been interviewed at least once.

The SHARE Research Data Center provides the SHARE survey data separately for each wave including all countries. Therefore, the SHARE data should be reduced to German respondents only. A detailed documentation of the SHARE data can be found [here][13].

2.1.2 SHARE GV-LINKAGE- OVERVIEW

A generated data file "*sharewX_rel7-1-0_gv_linkage*", the so-called 'overview' data set, includes only German respondents and was specifically generated for the project SHARE-RV.

This data set summarizes very useful [information][14] about the availability of the data, e.g. who consented and whether and which administrative data are available. It can be used in the early stages of analyses to find out for which respondents the administrative data are available.

Different from the standard SHARE modules, which are wave-specific, the 'overview' dataset includes all respondents who participated in at least one wave between Waves 3-7. This implies that not all cases included in the 'overview' dataset will merge when linking it to a specific wave of the SHARE data, e.g. merging the 'overview' data set with a German SHARE Wave 7 data set results in 3011 unmatched cases. These are respondents who did not participate in the German wave 7, such as the ones with the fictitious mergeid "DE-01", "DE-02" and "DE-03" (see [Figure 1][]).

The variables "share_w3", "share_w4", "share_w5", "share_w6", and "share_w7" indicate in which SHARE waves the respondent participated. The variables "share_w1", "share_w2" indicate whether a respondent participated in wave 1 and respectively in wave

---

[12] For detailed information see: www.share-eric.eu/home/overview.html

[13] http://www.share-eric.eu/data-documentation.html

[14] For detailed information see http://www.share-eric.eu/fileadmin/pdf_documentation/Wave_7_Updates___ Documents/SHARE-RV_Wave_7_Documents/English_SHARE-RV_7-1-0_Codebook_overview.pdf

2 or not under the condition that they have participated in at least one of the later waves (wave 3, 4, 5, 6, or 7). As the FDZ-RV provides two different data sources, the "Versicherungskontenstichprobe (VSKT)" and the "Versichertenrentenbestand (RTBN)", the variables "vskt" and "rtbn" summarize which data sets are available via FDZ-RV.

2.2 THE ADMINISTRATIVE RECORDS

2.2.1 VERSICHERTENKONTENSTICHPROBE (VSKT)

The data in the format of the VSKT is longitudinal and contains monthly information on respondents' employment histories from January of the year in which the respondent turned 14 until December of the year the respondent turned 65 (624 months in total). Each longitudinal variable is provided as a separate data set including 624 variables in wide-format which refer to one month each. The first variable (month 1) refers to the January the respondent turned 14, the second variable (month 2) the February the respondent turned 14 and so on. To assign the calendar time of each month, see Chapter 4.4. The suffix of each data set corresponds to the names of the longitudinal variables. The VSKT contains information e.g. on the type of employment, earnings points for each month, and child-care periods. The dataset 'fix' includes time-invariant information as demographics as well as accumulated pension entitlements. For detailed information please take a look at the VSKT codebook[15] or the VSKT User Information[16].

2.2.2 VERSICHERTENRENTENBESTAND (RTBN)

The RTBN is cross-sectional and available for retirees only. It summarizes respondents' benefits which are accumulated until retirement. In addition, the amount of the pension which is paid is included. For detailed information please take a look at to the RTBN codebook[17].

The two administrative data sets (VSKT and RTBN) are also not wave-specific but include all respondents who consented and whose administrative data are available. Again, not

---

[15] http://www.share-eric.eu/fileadmin/pdf_documentation/Wave_7_Updates___Documents/SHARE-RV_Wave_7_Documents/English_SHARE-RV_7-1-0_Codebook_VSKT.PDF

[16] http://www.share-eric.eu/fileadmin/pdf_documentation/Wave_7_Updates___Documents/SHARE-RV_Wave_7_Documents/SHARE-RV_7-1-0_VSKT-User-Information.pdf

[17] http://www.share-eric.eu/fileadmin/pdf_documentation/Wave_7_Updates___Documents/SHARE-RV_Wave_7_Documents/English_SHARE-RV_7-1-0_Codebook_RTBN.PDF

Stand: 15.04.2020

all cases will merge when combining the VSKT (the same holds for the RTBN) data set with any wave-specific SHARE data. The administrative data includes only those respondents whose data are available independently from the SHARE wave they participated in.

3. DATA ACCESS AND DOWNLOAD

Researchers must register for both datasets separately. Registered SHARE users will be informed about new releases and other important news automatically.

In order to access the SHARE data, researchers have to [register](#)[18]as SHARE users at the SHARE Research Data Center. Terms and conditions of the Research Data Center can be found there as well. Upon registration, users are able to download SHARE data. Once logged in, you can directly download the 'overview' dataset 'Linkage Release 7.1.0'. In order to get access to the administrative records of the FDZ-RV, researchers additionally have to register as users at the FDZ-RV. Please note: The FDZ-RV will only grant data access if the project leader has minimum a doctor's degree. The administrative data of the FDZ-RV which can be linked to the survey data will be provided on CD. Please follow the step-by-step instruction given in [Appendix A](#) to register as FDZ-RV user and request the administrative data. At the time of writing, the homepage is available in German only.

4. DATA HANDLING

4.1 DATA FORMATS AND LANGUAGES

All data is provided in Stata and SPSS format. For using other statistical programs, the data can be transferred by users themselves.

The datasets of the FDZ-RV are published in German. For English speaking users, a syntax for English labeling is provided via the Compact Disk of the FDZ-RV.

Please note that SHARE-Data and administrative RV-data are at least readable by Stata12 (characters in extended ASCII format).

---

[18] www.share-eric.eu/data-access/user-registration.html

Stand: 15.04.2020

If you use Stata 14 or newer, the German "Umlaut" will not be displayed correctly. In this case, either open the files by using the "useold"[19] command (should be installed from within Stata by typing "ssc install useold") or if this is not possible for you, use our provided German label file for Stata 14 or newer (UNICODE format).

Feel free to contact us, in case you have any other problems with the data formats!

4.2   HOW TO LINK THE DATA?

The administrative data are provided with the same identification number as used in the SHARE data (mergeid) so that it is linkable in the same way as all SHARE modules. easySHARE[20] can be linked with the administrative data using the same identifier (mergeid) as well.

Although the project SHARE-RV started in the third wave of SHARE (SHARELIFE), the linkage is not limited to data of this wave. As all SHARE respondents have a stable identifier, the administrative data can also be linked with data of the same respondents of previous and following waves. As a consequence, the administrative data sets consist of records of respondents who may have participated in different SHARE waves.

SHARE's German target population consists of all persons aged 50 years and over at the time of sampling who have their regular domicile in Germany. In addition, the respondents' current partners living in the same household will also be interviewed, regardless of their age. Therefore, it is possible to do analyses not only on individual level but also on household level if both took part in the SHARE interview. In order to identify persons belonging to the same household please refer to the specific-wave household identifier (hhid*). More information could be found in the SHARE User Guide[21], Chapter 'Merging the data'.

4.3   REPORTING YEAR

The variable 'reporting year' in the administrative data sets gives the user information about the year the data refers to. Given that it was not always possible to update the

---

[19] https://ideas.repec.org/c/boc/bocode/s458054.html

[20] www.share-eric.eu/special-data-sets/easyshare.html

[21] www.share-eric.eu/fileadmin/pdf_documentation/SHARE_release_guide_7-1-0.pdf

Stand: 15.04.2020

data for all respondents every year, the data set includes the most recent version available for each respondent. This means, for example, that the administrative data of the longitudinal data set includes information until in December 2016 if the reporting year is 2016.

4.4 HOW CAN I FIND OUT IN WHICH YEAR AND MONTH AN EVENT HAPPENED BASED ON VSKT?

Each longitudinal variable is in the same format (wide), consisting of 624 variables which refer to 624 months of respondents' biographies. These months always correspond to the same period in respondents' life: beginning in January of the year the respondent becomes 14 years old (`var'_1) until December of the year the respondent becomes 65 years old (`var'_624).

An example: Respondent Anna is born in March 1950. Therefore, the variable SES_1 refers to January 1964 since she turned 14 that year. She started her first job in month 77 (SES_77=13 for the first time → start=77). To find out the 'real' year and month of that event, the following formula should be used:

$$\text{Year} = \text{year of birth} + 14 + \text{int}^{[22]}[(\text{start}-1)/12]$$
$$\text{Month} = \text{start} - (\text{Year} - \text{year of birth} -14) *12$$

For our example this is:

Year = 1950 +14 + int [(77-1)/12]

= 1964 + 6

= 1970

Month = 77 – (1970-1950-14)*12

= 5

We now know that the event-month 77 refers to May of 1970.

---

[22] In SPSS the command is ‚truncate‘, in STATA ‚int‘ or 'trunc'

Stand: 15.04.2020

4.5 WHAT ARE THE EARNING POINTS EXACTLY?

Basically, earning points are based on the level of the income in each individual year. Thereafter, voluntary contributions will be added after they have been converted into remuneration as well as earnings points with a fixed value (e.g., for child-raising periods). For each year, individuals' income will be related to the average income of all insured persons. If the income corresponds to this average income, individuals will receive one point. If the income was less, they will receive fewer points and if the income was higher, they will receive more points. Income from employment in the new Federal States (East Germany) will be updated to the level of the old Federal States by using a predetermined factor. These points are called earning points (Deutsche Rentenversicherung Bund, 2012). For more information please refer to the Pension Law[23].

---

[23] https://www.deutsche-rentenversicherung.de/SharedDocs/Downloads/DE/Broschueren/international/europ_ vereinbarungen_fremdsprachig/leben_und_arbeiten_europa_englisch.pdf?__blob=publicationFile&v=4

PLEASE NOTE: When using SHARE-RV data for publications, please cite the dataset as "Forschungsdatenzentrum der Rentenversicherung, Max-Planck-Institut für Sozialrecht und Sozialpolitik (2020). SHARE-RV. Release version: 7.1.0. SHARE-ERIC. Dataset. DOI: 10.6103/SHARE.SHARE-RV.710" and add the obligatory basic disclaimer for SHARE waves 1-7 which can be found [here][24].

When using SHARE-RV data, users are additionally required to cite the following publication:

Börsch-Supan, A., C. Czaplicki, S. Friedel, I. Herold, J. Korbmacher, T. Mika (2020). *SHARE-RV: Linked Data to Study Aging in Germany.* In: Winker, P., T. v. Büttner, R. Riphahn, W. Smolny, J. Wagner (eds.): Jahrbücher für Nationalökonomie und Statistik, *240*(1), 121-132. Berlin: De Gruyter. DOI: 10.1515/jbnst-2018-0034.

In addition, please do not forget to provide us with information about all publications using SHARE data by emailing to info@share-project.org.

Responsible for preparation of the SHARE-RV data and the documentation is:

Imke Herold

Version: 7-1-0 (31.03.2020)
Contact: info@share-project.org with "SHARE-RV" in the subject

---

[24] http://www.share-eric.eu/data-access/citation-requirements.html

Stand: 15.04.2020

APPENDIX A

Please note: The FDZ-RV will only grant data access if the project leader has minimum a doctor's degree.

1. Go to the FDZ-RV website and click on "Nutzungsantrag" > "zum Erstantrag"

2. Click on "Neuer Antrag" to get the form which has to be filled out.

3. Please provide the requested information on the next two screens and click on "Weiter".

1. By choosing "Scientific Use File" and then "Themenfile", it is possible to select "SHARE-RV". Then click on "Weiter" and check your personal data carefully.

2. In order to send your application, **click on "Antrag senden"**.

    Please print and sign the document and send it to the FDZ-RV.

    Either by post: Deutsche Rentenversicherung Bund
                          FDZ-RV – GB0600
                          10704 Berlin

    Or as pdf by email: FDZ-RV-Berlin@drv-bund.de with "Antrag auf Datennutzung" in the subject.

6. After you have sent your application, it will be reviewed by the FDZ-RV. After a successful review, you will receive a User Contract by mail. Here are a few essential regulations of the contract:

    *   The usage of SHARE-RV data is limited to those persons, who are mentioned by name in the contract. It is, furthermore, not allowed to publish the data.
    *   The usage of SHARE-RV data is limited to the duration of the contract. The data users are obligated to delete all data files by end of the contract.

7. The user contract has to be signed by the applicant, the head of the institution as well as the data administrator and sent back to the FDZ-RV. After the FDZ-RV received the signed contract, a CD with the data will be provided for free. **Please note that the entire order process takes at least 2-3 weeks.**

Stand: 15.04.2020