

Rules for visiting researchers at the RDSC

Principles and rules

Technical Report 2021-02, Version: 1-0

Valid from 01 March 2021

DOI: 10.12757/75721.1

Deutsche Bundesbank, Research Data and Service Centre

Keywords: visiting researcher rules, disclosure control, output control, publication control, data protection

Version: 1-0

DOI: 10.12757/75721.1

Citation: Research Data and Service Centre (2021). Rules for visiting researchers at the RDSC, Technical Report 2021-02 - Version 1-0, Deutsche Bundesbank, Research Data and Service Centre.

Contents

1 Introduction	4
2 Rules on working in the secure environment of the RDSC	5
3 Principles of output control	7
Principle O.1: Anonymity	7
Principle O.2: Verifiability	9
Principle O.3: Reproducibility	9
Principle O.4: Economical use of resources	10
Principle O.5: Responsibility	11
4 Principles of publication control	13
Principle P.1: Review of all publications by the RDSC	13
Principle P.2: Criteria for the publication of research outputs	13
5 Glossary	14

1 Introduction

The Research Data and Service Centre (RDSC) offers high-quality services for accessing microdata held by the Deutsche Bundesbank. Simple and quick application channels, excellent user support and data enrichment are the RDSC's core business. In this context, the data provided to the RDSC are handled responsibly and in compliance with all legal and internal regulations on data processing and data disclosure.

Visiting researchers working with Bundesbank microdata are responsible for ensuring that their calculation results do not violate data confidentiality requirements. If their research outputs leave the RDSC's secure environment, they must ensure that their calculation results do not contain any data that can be traced back to individual observation units or statistical units such as banks, (non-financial) corporations, individuals or households.

Compliance with these data confidentiality requirements when producing these results is additionally checked by RDSC staff¹. If these requirements are not fulfilled, the calculation results cannot be released and/or publication cannot be approved. In the event of particularly serious breaches, additional measures specified in the visiting researcher's contract may be taken.

This document is intended to help visiting researchers more easily comply with the requirements stipulated for their time at the Bundesbank, the release of their calculation results, and the approval of publications based on these results. Furthermore, these principles and rules also boost confidence among data providers, such as banks and enterprises, and data producers, such as Bundesbank business units, that their data are being handled securely.

Visiting researchers can contact the RDSC at any time at fdsz-data@bundesbank.de if anything is unclear or if they have questions. The RDSC strongly encourages researchers to clarify any points that are unclear prior to their first visit at the RDSC.

The RDSC reserves the right to amend, supplement or expand on the following principles and rules, if necessary.

This document contains principles and rules for (i) working in the secure environment of the RDSC, (ii) output control, and (iii) publication control. In addition, the final chapter includes a Glossary featuring many definitions of the terms used in this document.

¹ Under normal circumstances, it is therefore the RDSC that carries out these checks. In individual cases, other business units of the Bundesbank may be responsible for performing them.

2 Rules on working in the secure environment of the RDSC

When dealing with microdata as part of a scientific research project, visiting researchers are obliged to maintain data secrecy. In order to fulfil the criteria of *de facto anonymity*²⁾ in the secure environment of the RDSC, all visiting researchers must comply with the following rules:

1. **Access to data only with authorisation.** Within the scope of a research project, data may only be used by authorised visiting researchers. Authorised visiting researchers have signed a contract with the RDSC and enter into a set of obligations upon commencing their first research visit. Microdata may not be disclosed to third parties. Co-authors who are not authorised visiting researchers are also considered to be third parties.
2. **No use of external hardware.** The use of laptops, smartphones, tablets and other electronic devices at visiting researcher workstations is prohibited.
3. **No copies of microdata.** All data must remain in the project's designated storage location. Data may not be copied, printed, photographed or saved outside of the project folder (e.g. on the local hard drive or desktop).
4. **No re-identification.** When using microdata, re-identification is strictly prohibited. Visiting researchers may not de-anonymise anonymised microdata.
5. **Adherence to specified folder structure.** Visiting researchers are given access to a personal project folder labelled with their project number. The directory structure within the project folder is intended to keep original data (`.\data\orig`), any external data used (`.\data\ext`), necessary temporary data (`.\data\temp`), program codes (`.\code`), results (`.\results`), and results for output control (`.\transfer`) separate from each other. This folder structure (subfolders included) specified by the RDSC may not be altered or deleted. Visiting researchers may create their own folders within their personal project folders according to the specified rules. The maximum path length for file names including the corresponding path structure is 130 characters.
6. **Usage rights for external data.** In order to use external (i.e. non-Bundesbank) data as part of a research project, confirmation of data usage rights must be submitted in writing to the RDSC before the start of the project.
7. **Timely submission of external data.** External data intended for use at a visiting researcher workstation must be submitted by email to the RDSC for review at least five working days before the scheduled start of the research visit.
8. **Timely submission of external program codes.** External program codes intended for use at a visiting researcher workstation must be submitted by email to the RDSC or another responsible business unit of the Bundesbank. To ensure that there is sufficient time for review, external program codes must be submitted at least five working days before the scheduled start of the research visit.
9. **Use of approved software.** Stata is available as standard at visiting researcher workstations. Use of Matlab, Python and R is possible in principle, but must be agreed upon with the RDSC in writing in advance. Calculation results produced with Microsoft Excel or saved in an Excel file will not be accepted for output control or release.
10. **No modifications to provided hardware or software.** Unauthorised system modifications, such as deactivating protective functions or setting up local administration rights for PC workstations in the Bundesbank network, are prohibited. Autonomously installing software on Bundesbank devices and using non-approved/private software (including portable versions) on

2 Act on Statistics for Federal Purposes (*Bundesstatistikgesetz – BStatG*), Section 16(6) number 1: “[...] individual data [where] attributing the individual data to the relevant respondents or persons concerned requires unreasonable effort in terms of time, cost and manpower (de facto anonymised individual data)”.

Bundesbank devices is prohibited.

11. **Adherence to project duration.** The data provided may only be used for the duration of the project. For visiting researchers, the project ends at the same time as their contract.

3 Principles of output control

The following five principles and associated rules are intended to help visiting researchers more easily comply with the rules for checking research outputs with regard to data confidentiality (output control).

Principle O.1: Anonymity

All calculation results that leave the secure environment of the RDSC must be *absolutely anonymised*. Results which are absolutely anonymised cannot be identified directly through labelling of reporting agents or other legal or natural persons, legal entities or branches (e.g. by means of their name or address or an officially issued identification code) or indirectly through deduction. Consideration should be given to all possible ways and means that a third party could, in appropriate form, employ to identify the microdata. Visiting researchers are responsible for ensuring that their calculation results fulfil these criteria. In order to ensure complete anonymity, compliance with the following rules is required:

1. **No release of identifiers.** Identifiers may not be included in either the calculation results or the program codes.
2. **No release of microdata.** Research outputs may not include any microdata. This applies to the release of datasets, tables and charts as well as to commands and comments in program codes (e.g. Stata do files) and log files (e.g. Stata log files). Minima and maxima are generally considered to be microdata. Quantiles are also considered to be microdata if they do not fulfil the criteria of *rule 5: calculation of quantiles*.
3. **Adherence to minimum sample size.** Results intended for release must be based on at least five different observation units.
4. **Adherence to p% or dominance rule.** Even if the reported results contain five or more observations, major observation units could still be indirectly identified if their combined share of the value under analysis is large enough. Therefore, the combined shares of the two largest observation units should not exceed 85% of the total value under analysis.³⁾
5. **Calculation of quantiles.** If quantiles q are calculated for a population or subgroup, there must be at least five different observation units in the interval between the largest quantile and the maximum. For reasons of symmetry, this is also required for the interval between the minimum and the smallest quantile. The dominance criterion must likewise be fulfilled. If there are fewer than five different observation units in the interval from quantile q_p to quantile $q_{(p+1)}$, at least 15% of the value of the relevant variables must lie outside the interval for at least one of the observation units.
6. **Confidentiality requirement for multiple tables.** If calculation results are produced based initially on the population G but subsequently on a subset $X \subset G$, compliance with the rules mentioned above must also be verified for the difference. If this is not done, individual observation units could potentially be identified on the basis of this difference.
7. **Treatment of dummy variables.** If the mean value of a dummy variable coded as 0 and 1 is produced, at least five different observation units must have a value of 0 and at least five different observation units must have a value of 1.

3 This refers to non-transformed values.

8. **Treatment of zeros and missing values.** In the case of dichotomous and categorical variables, zeros are permissible values in descriptive statistics and regression analyses provided that they are not used as a code for missing values. In the case of constant variables, zeros may be used in regression analyses. In descriptive statistics, zeros and missing values may not be taken into account when determining the number of different observation units.
9. **Treatment of regional information.** Descriptive analyses at regional level are not permitted if the regional units are smaller than administrative districts (NUTS3). In cases of analyses at regional level or for specific regional units, the results must be based on at least 20 observation units and the dominance criterion must be fulfilled. For analyses at regional level using household data, such as the Panel on Household Finances (PHF) or the Bundesbank Online Panel Households (BOP-HH), a minimum sample size of five observation units is sufficient and the dominance criterion must be fulfilled. In regressions, regional dummies may also be used for smaller regional units, for example to control for region-specific effects. Please note that the coefficients for regional dummies are generally not released. Furthermore, regional information must not allow the observed regional unit to be identified indirectly.

A separate set of confidentiality rules applies to certain microdata. For this reason, the following additional criteria need to be fulfilled when working with such data:

- i. **External position of banks (AUSTA – “Auslandsstatus”).** A bank (parent) can maintain multiple legally dependent branches and/or legally independent subsidiaries. Therefore, when conducting analyses at the level of foreign affiliates, the number of associated parents must always be checked.
- ii. **Microdatabase Direct investment (MiDi).** When determining sample sizes, it must be ensured that an enterprise is counted only once and not multiple times, even if it makes more than one investment. Superordinated enterprises must likewise be considered entities subject to protection. In the case of outward investment, these are the *nu2* (the investment enterprise abroad / foreign subsidiary), the *num* (German reporting agent) and, if applicable, the *nui* (German parent of the reporting agent) where multiple *nums* belong to one *nui* as one unit. In the case of inward investment, the following are to be protected accordingly: *num*, *nu4* (foreign parent), *noa* (group parent of the foreign parent, if applicable).
- iii. **Credit register for loans of €1 million or more.** The minimum sample size must be adhered to for both borrowers and lenders. If there is information on superordinated lenders (lender group) and borrowers (borrower unit), these must likewise be considered entities subject to protection.
- iv. **Securities Holdings Statistics (SHS-Base plus).** The entity subject to protection is the reporting institution and not the security. Reporting institutions’ holdings of their own securities, i.e. a combination of depositor sectors 1221/1222/1223, ISIN, and bank, are problematic, as there is a possibility of ascertaining the securities issuer, and thereby the reporting institution, on the basis of the ISIN. For this reason, it is also not permitted to provide names of securities or other information that would enable the issuer to be identified.
- v. **Statistics on International Trade in Services (SITS).** The SITS dataset is concerned with services transactions, i.e. each row depicts a transaction reported for a given date (year and month). Because an enterprise will usually carry out more than one transaction, it is possible for multiple rows to contain observations for one and the same enterprise. The number of rows showing one and the same month therefore does not correspond to the number of enterprises. When determining the sample size, it must be ensured that an enterprise is counted only once

and not multiple times, even if it carries out more than one transaction.

- vi. **Corporate balance sheets from refinancing (Ustan).** Some enterprises report multiple sets of financial statements each year. Therefore, when determining the sample size, it must be ensured that these enterprises are counted only once and not multiple times.

Principle O.2: Verifiability

In order to verify compliance with the anonymity requirements specified in *principle O.1*, it must be possible for RDSC staff to verify the production of research outputs within a reasonable amount of time. For this reason, visiting researchers are required to comply with the following rules:

1. **Master file.** A master file must be created that contains all relevant information on the research project and calls all of the subprograms used. Stata templates ("*0_master.do*", "*1_data_preparation.do*", "*2_descriptive_analysis.do*" and "*3_regressions.do*") are provided by the RDSC.
2. **Logging.** The log function (log file) must be activated for every program code. Logging must commence before the first statement and before the description of the research project's content.
3. **Order and structure within program codes.** Program codes must be structured in a visually clear way so that individual program blocks (program header, individual analytical steps, etc.) are visually distinct. Loops must be indented. Long programs or analytical steps must be divided into smaller program files, e.g. data preparation, descriptive analysis, regressions (see Stata templates under *rule 2: logging*).
4. **Comments in program codes.** The program code must feature sufficient commentary so that even persons who are unfamiliar with the project are able to understand it within a reasonable amount of time.
5. **Clear naming of output files.** The names of all output files must begin with the same name as the program used to generate the file and must be numbered in a logical manner.
6. **Clear naming of variables.** All of the names assigned must be as informative as possible and used consistently. Variable labels and short descriptions of the variables must be provided for all data generated by the user themselves as well as for all data sourced externally. If (categorical) variables are newly created or modified, corresponding value labels need to be assigned to these values.
7. **Specification of sample size.** The number of underlying observation units for the calculations must be specified for each format used to present the results (e.g. tables, charts, analyses). In Stata, the sample sizes can be produced and checked using the *nobsdes5* and *nobsreg5* commands provided by the RDSC. The corresponding ado files are available in the RDSC's *ado_library* and on the RDSC's website.
8. **Re-examination.** In the event of minor changes to the program codes and resubmission of the output, the changes made must be indicated explicitly. Where possible, visiting researchers should submit only the affected program elements for re-examination.

Principle O.3: Reproducibility

In order to verify compliance with the anonymity requirements specified in *principle O.1*, all calculation results submitted for review must be reproducible. Visiting researchers are required to

comply with the following rules:

1. **Reproducibility of program codes.** All calculation results must be generated seamlessly by a program named "*O_master.do*" that can be executed without error and that must contain all of the analysis programs used over the course of the project. This program must begin by loading the original data provided by the RDSC. The program must always execute the same processing steps and produce exactly the same results as those already submitted for review. In the "*O_master.do*" file, every program call must be followed by a short description of the content of the relevant subprogram.
2. **Reproducibility of software.** All software used to generate the calculation results must be clearly stated at the start of the "*O_master.do*" file (name and version number). Alongside the version number of the analysis software, this also includes the names of all packages (e.g. R, Python, Matlab) used in addition to the software or any program codes generated by other users (e.g. Stata ado files), and their respective version numbers. This also applies to software provided by the RDSC.
3. **Reproducibility of data.** All RDSC datasets used to generate the calculation results must be clearly and unambiguously stated at the start of the "*O_master.do*" file (DOI if available, statistics and year). Any datasets used that originate from external data providers must be submitted, together with appropriate documentation on the data and the sources, before they are used in the research project.
4. **Reproducibility of publications.** All published calculation results must be able to be found quickly, easily and clearly in the output produced by the analysis programs. For this purpose, there must be a clear relationship between the analysis program, the output and the publication. This can be achieved, for example, if the researcher attaches a document to their publication that links each calculation result to a table name in the output. All calculation results featured in the text of this document must also be listed alongside any tables and charts. The RDSC recommends creating a "*master_yyyymmdd_publication.do*" file containing all of the calculation results with references to the corresponding calculations, tables and charts in the publication. If the publication is determined to be non-reproducible, it cannot be approved for release.

Principle O.4: Economical use of resources

As a general rule, results that leave the RDSC are meant to be for the purposes of a publication. For this reason, visiting researchers should be guided by the principle of using resources economically when deciding which results to submit for review. The number of calculation results presented should be in line with what would normally be expected within the scope of an empirical scientific article. In addition, when choosing which results to submit for review, visiting researchers should bear in mind that, in particular, the checks relating to the "confidentiality requirement for multiple tables" (see *principle O.1 – rule 6*) are very time-consuming. Generally, visiting researchers must heed the following rules:

1. **No exploratory analyses.** Only analyses that could potentially be published may be submitted for review. The task of sifting outputs worthy of publication from the large number of calculation results often generated in the context of exploratory analyses or when testing different model specifications for regressions should be completed during the research visit at the RDSC.

2. **Maximum number of lines for output.** Under normal circumstances, the total number of lines submitted for review in the course of a research project must not exceed a maximum of 2,500. Charts to be approved are generally treated on a flat-rate basis as equivalent to 25 lines. Since pure code files (e.g. Stata do files) are not permitted to contain any results, these are not included in the line count. The RDSC notifies visiting researchers as soon as a threshold of 2,000 lines has been exceeded.
3. **Increase in the maximum number of lines for output.** In extreme exceptional cases, visiting researchers may request a one-off increase in the maximum number of lines permitted. To do so, they must provide a comprehensive explanation of why they need additional output. Visiting researchers will need to provide detail on:
 - a. why the maximum number of lines is insufficient;
 - b. how they intend to satisfy the confidentiality requirement for multiple tables (see *principle O.1 – rule 6*);
 - c. how many additional lines are needed up to completion of the project;
 - d. why these additional lines are needed up to project completion (detailed description of planned activities with estimated quantities);
 - e. detailed explanation of project-specific measures taken to ensure compliance with the new limit.

On the basis of this justification, the RDSC establishes a new number of lines for the output. If the grounds are inadequate, the request for an increase in the maximum output line count will be turned down, with the RDSC providing a justification for its decision. A rejected application may be amended and resubmitted once.
4. **Maximum number of lines for output in projects involving researchers of the Deutsche Bundesbank.** Output is not counted if researchers of the Deutsche Bundesbank carry out the output control themselves. This exception does not apply to projects where the RDSC carries out or is required to carry out the output control.
5. **(IT) resource-conscious programming.** Visiting researchers should take a parsimonious approach when deciding which calculations of a more extensive nature are truly necessary for the research project. This applies to runtimes, system load and the data volume generated.

Principle O.5: Responsibility

Visiting researchers are responsible for ensuring compliance with all of the principles and rules set out in this document. This is covered in detail as part of the set of obligations that all visiting researchers are required to enter into before embarking upon their research project. Failure to comply with the rules will result in the RDSC refusing to release the calculation results. Visiting researchers must heed the following rules:

1. **Checking of all calculation results for release.** Before visiting researchers can use calculation results outside the secure environment, the RDSC must first check whether the visiting researchers have applied the principles of output control to them, and the calculation results must be approved for release by the RDSC. To that end, visiting researchers must inform the RDSC by email whether calculation results require checking and which those are. Results for review should always be saved in a subfolder of “*transfer*”, with the relevant date as the file name. Additional information that is necessary as part of the review but does not need to be published must be filed in a subfolder of “*results*” named with the same date. Calcula-

tion results may not be sent in the form of an email attachment. The RDSC does not perform plausibility checks on the content of the research outputs it receives.

2. **Functioning program codes.** If the program code contains syntax or other errors, these will be left uncorrected by the RDSC and the visiting researchers will be asked to correct their code.
3. **Format for output control.** Results and program codes are only accepted for output control if they are editable and are submitted as plain text files. Charts must be in a read-only format.

4 Principles of publication control

The following two principles and associated rules are intended to help visiting researchers more easily comply with the rules for checking publications (“publication control”).

Principle P.1: Review of all publications by the RDSC

1. **Review of all publications by the RDSC.** Research outputs containing calculation results may not be published until the checker has given written approval. Approval may be withheld if the results intended for publication do not comply with the criteria specified in *principle P.2*. The review process for publications may take up to 10 working days.
2. **Obligation to submit.** It is the responsibility of visiting researchers to submit to the RDSC all publications they are preparing that contain research outputs from analyses carried out during their time at the RDSC.
3. **Reference to previously approved calculation results / reproducibility of published results.** Publications may only contain calculation results approved by the RDSC. When submitting a publication for review, visiting researchers must comply with the requirements set out under *principle O.3, rule: reproducibility of publications*. If they do not, the publication cannot be approved for release.
4. **Multiple publications.** Further publications based solely on previously approved publications do not require authoriser approval. Calculation results that have not already been included in a previously approved publication must be marked up and the publication must be submitted to the RDSC for review.

Principle P.2: Criteria for the publication of research outputs

A publication must also meet the following criteria:

1. **Specification of research project.** Each publication must state the project number.
2. **Declaration of the type of data access.** Each publication must specify the type of data access, i.e. research visit, controlled remote execution (CRE) or scientific use file (SUF).
3. **Specification of the datasets used.** All datasets used in the research project must be cited, giving the name and, if available, the DOI and, in addition, the associated data documentation or corresponding publications.
4. **Citation.** All charts and tables must appear with a citation: *Source: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, <name of the microdataset used from the standard RDSC list (where appropriate with the common abbreviation)>, <period during which the microdata were used>, own calculations.*

5 Glossary

Term	Explanation
Absolute anonymity	Results which are absolutely anonymised cannot be identified directly through labelling of reporting agents or other legal or natural persons, legal entities or branches (e.g. by means of their name or address or an officially issued identification code) or indirectly through deduction. Consideration should be given to all possible ways and means that a third party could reasonably employ to identify the microdata.
Aggregate data	Aggregate data combine several microdata into one single value.
Calculation results	Calculation results are all works produced as part of the research project, particularly project outlines, data analyses, academic papers and presentations, including those still in draft form, as well as all research findings and knowledge gained over the duration of the research project.
Checker	The RDSC or, in individual cases, another Bundesbank business unit responsible for reviewing compliance with data confidentiality requirements.
Controlled remote execution (CRE)	The RDSC offers controlled remote execution for some of the microdatasets available. CRE allows researchers to analyse the data without being physically present on RDSC premises. This is done by sending their data analysis programs to the RDSC.
Data confidentiality	The principle of integrity and confidentiality (Article 5(1) point (f) of the GDPR) requires that personal data be processed in a manner that ensures appropriate security of those data. In particular, arrangements must be made to ensure that the personal data are protected: <ol style="list-style-type: none"> 1. against unauthorised and unlawful processing (confidentiality); 2. against accidental loss, destruction or damage (integrity). In order to achieve this, appropriate technical and organisational measures (TOMs) must be put in place.
Data documentation	The RDSC provides data documentation including metadata for almost all of the microdatasets available. These can be downloaded from the RDSC's website.
Data secrecy	Individuals involved in data processing may not process personal data without authorisation (data secrecy). They shall be obligated to maintain data secrecy when taking up their duties. The obligation to maintain data secrecy shall continue even once they are no longer actively performing those activities (see Section 53 of the Federal Data Protection Act (<i>Bundesdatenschutzgesetz</i>)).
De facto anonymity	Information is classed as de facto anonymous if attributing the underlying microdata to the relevant entity would require unreasonable effort in terms of time, cost and manpower. The formally anonymised data provided by the RDSC in the secure environment are also considered de facto anonymised (see Section 16(6) number 2 of the Act on Statistics for Federal Purposes (<i>Bundesstatistikgesetz – BStatG</i>)).

...

		...
De-anonymisation		De-anonymisation is the deliberate reversal of previously performed data anonymisation to enable re-identification of the underlying microdata.
Descriptive analysis		By means of descriptive analysis, empirical data are described and clearly presented in the form of a table, metric or charts.
Digital object identifier (DOI)		The digital object identifier (DOI) as defined by ISO 26324 is a unique, persistent digital identifier for physical, digital or abstract objects. All standard datasets provided by the RDSC are registered with a DOI.
Direct identification		Identification of reporting agents or other legal or natural persons, entities or branches by name, address or any officially issued identification code, thereby revealing individual data.
Dominance criterion		Observation units differ in size. Even if a value based on five or more observation units is reported, there is the risk of major observation units being identified indirectly if together they contribute a sufficiently large share to that value (dominance criterion). The possibility of indirectly identifying observation units in this way must be eliminated. It must therefore be shown that the observation units used as a basis for calculating a value do not include two observation units that are significantly larger than the remaining observation units. As a rule of thumb, the two largest observation units should not account for more than 85% of the total value when combined.
Entity subject to protection		See definition of <i>observation unit</i> .
Formal anonymity		Formal anonymity is achieved by removing the direct identifiers from the data material but leaving other characteristics untouched.
Identifier		An identifier is a tag serving to uniquely identify a specific object.
Indirect identification		Identification of reporting agents or other legal or natural persons, entities or branches by deduction, thereby revealing individual data.
Microdata		Microdata are highly granular data that directly refer to units of observation.
Minimum sample size		This document sets out a minimum sample size of five observation units as standard. See definition of <i>sample size</i> .
Obligations		On the day of their first research visit, visiting researchers are briefed by RDSC staff on the obligation pursuant to Section 1(1) of the Act on the Formal Obligation of Persons Without Civil Servant Status (<i>Verpflichtungsgesetz</i>) of 2 March 1974 and their obligation to maintain data secrecy.
Observation units		Observation units are the entities subject to protection, i.e. all reporting agents or other legal or natural persons, entities or branches. These could be banks, enterprises, households, borrowers or lenders, for example.
Output		See definition of <i>calculation results</i> .
Output control		The RDSC checks that visiting researchers have complied with the principles and guidelines set out in <i>Section 3</i> of this document.
		...

	...
Plain text files	Plain text refers to data that can be converted directly into text using character encoding. Examples of this are .csv or .txt files.
Project number	The RDSC assigns an eight-digit project number to each research project at the outset. This consists of the year (YYYY) and a consecutive number string (xxxx): YYYY/xxxx.
Project folder	Every research project involving at least one research visit has its own dedicated project folder on the Bundesbank's internal drive. This is labelled with the project number and is accessible to visiting researchers during their research visit.
Publication	<p>The term publication refers, amongst other things, to the presentation of research outputs:</p> <ul style="list-style-type: none">– in journal articles;– in papers within edited volumes;– in books;– in working paper series and discussion papers published by institutions or enterprises;– in dissertations;– in online publications, especially on academic platforms such as SSRN.com and RePEc.org;– on conference websites or personal websites such as Google Pages, etc.;– on the staff webpages of academic institutions;– in conference submissions in the form of working papers;– in generally accessible presentation slides;– in generally accessible Bachelor and Master theses. <p>A presentation is deemed generally accessible if it is technically suitable and intended to provide the public, i.e. a non-individually determinable group of persons, with information. Publication within the meaning of subsection 1 excludes the presentation of research outputs, provided they are not made available online:</p> <ul style="list-style-type: none">– in oral presentations;– in presentation slides;– in conference submissions (e.g. summaries, abstracts);– in Bachelor and Master theses. <p>Presenting research outputs in these ways does not require the prior written consent of the RDSC or another Bundesbank business unit responsible for this in individual cases; however, the confidentiality and data protection requirements pursuant to these terms and conditions shall apply in these cases as they do to all forms of public presentation.</p>
Publication control	The RDSC checks that visiting researchers have complied with the principles and guidelines set out in <i>Section 4</i> of this document.
Re-identification	See definition of <i>de-anonymisation</i> .
Research outputs	See definition of <i>calculation results</i> .
	...

...

Research visit	The period of time that visiting researchers spend physically present on RDSC premises. Visiting researchers must work at a visiting researcher workstation in the secure environment.
Sample size	Number of observation units an analysis is based on. This refers to the different IDs and not the total number of observations.
Secure environment	In the secure environment, visiting researchers can work on a PC provided by the RDSC that has neither an internet connection nor print functionality. The use of personal electronic devices such as mobile phones, laptops or tablets in the secure environment is strictly prohibited. Results which have not been reviewed may not leave the secure environment.
Scientific use files	Scientific use files (SUF) are standardised datasets compiled by the RDSC for selected statistics. As the data are, in this case, de facto anonymised, they may be used outside the secure environment. The potential for analysis offered by SUFs is inferior to that provided by a research visit, but they are designed to be suitable for a large proportion of scientific research projects.
Third parties	Third parties are all persons who are not visiting researchers or employees of the RDSC. This includes co-authors without direct data access.
Visiting researchers	All visiting researchers sign a contract with the RDSC and enter into a set of obligations upon their first research visit. Within the context of their research project, visiting researchers have direct access to confidential microdata held by the Bundesbank.
Visiting researcher workstation	The workstation where visiting researchers can work in the RDSC secure environment.
