DEUTSCHE
BUNDESBANK
EUROSYSTEM

Data Access to Micro Data of the Deutsche Bundesbank

Technical Report 2019-02

Tobias Schönberg (Research Data and Service Centre)

Disclaimer

The views expressed in this technical report are personal views and do not necessarily reflect the views of the Deutsche Bundesbank.

Citation

Schönberg, T (2019). Data Access to Micro Data of the Deutsche Bundesbank. Technical Report 2019-02, Deutsche Bundesbank, Research Data and Service Centre.

Abstract

This technical report provides an overview of the micro data access process at the Deutsche Bundesbank. Micro data enable data users to observe (statistical) behaviour of single statistical units. They are of major importance for many research fields in economics like financial stability analysis or banking supervision. The report describes different levels of micro data anonymity and user groups of these data at the Bundesbank. It also gives an overview of the IMIDIAS initiative – an internal initiative to support data sharing within the bank - and the five safes framework for data access at the Research Data and Service Centre of the Bundesbank.

Acknowledgements

I gratefully acknowledge the constructive comments and contributions from Stefan Bender, Martin Eisele, Christian Hirsch and Felix Sinn. Thankfully, Felix Sinn provided the text for chapter 4.

Content

1. Introduction

This technical report provides an overview of the micro data access process at the Deutsche Bundesbank. *Micro data* in the context of this document are data collected at the individual level and contain personal or factual information on statistical units.[1]

Chapter 2 lays out the distinction between micro, macro and granular data and introduces three different user groups of micro data at the Bundesbank. It also provides an overview of different degrees of micro data anonymity. Chapter 3 describes the IMIDIAS initiative which is a Bundesbank initiative with the goal to provide an integrated, cross-divisional information system for analysis and research purposes. Chapter 4 gives an overview on the legal framework for data access of various Bundesbank micro datasets.

Chapters 5 and 6 show how data access is practically implemented at the Bundesbank, whereby Chapter 5 focusses on internal analysts and Chapter 6 on internal and external researchers. A unit called *Internal Service for Micro Data Analysis* handles data access requests of internal analysts following a multi-level approach (modelled after the ESCB standard approach). The *Research Data and Service Centre* (RDSC) processes requests of internal and external researchers closely following the Five Safes framework, initially developed by Felix Ritchie at the UK Office for National Statistics. Chapter 7 concludes.

2. Definitions

This paper lays out the legal, organisational and technical framework for data access to micro data of the Deutsche Bundesbank. Before going into detail on the data access conditions, it is important to get an understanding of the term micro data, the different user groups of micro data at the Bundesbank and different degrees of micro data anonymity. The following chapters provide an overview of these topics.

2.1 A distinction between micro, macro and granular data

As the term *micro data* is widely used in many different contexts, it is easy to make the assumption that everyone has the same understanding of its meaning. However, sometimes this meaning is not clear. Therefore, in order to avoid confusion, the following paragraphs provide a definition of micro data and shed some light on the relationship between micro, macro and granular data (see Figure 1).

Micro data in the context of this document means data which contain personal or factual information on statistical units.[2] They are collected at the individual level (e.g. individual enti-

[1] See: Economic Commission for Europe of the United Nations, (2000), p.34.

ties, transactions or instruments). Most of the time with central banking statistics collected at the individual level, these micro data are not anonymised in their "raw" form. They enable data users to observe (statistical) behaviour of single statistical units.

In contrast, **macro data** is defined as observation data gained by a purposeful aggregation of statistical micro data.[3] With macro data, the focus is not at the individual level. For example, they are used to observe (statistical) behaviour of groups or the economy as a whole.

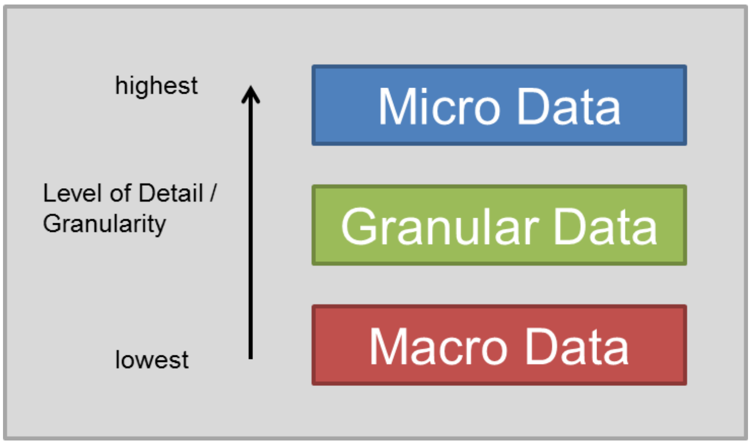

Figure 1: Relationship between micro, macro and granular data (own illustration).

In the literal sense, granularity refers to the level of detail of data: the greater the granularity, the deeper the level of detail. It is therefore not surprising that the term **granular data** is often used as a synonym for micro data. However, following Tissot (2015), it is actually a kind of intermediate data category between micro and macro data. Granular data can be less detailed than micro data, e.g. due to data confidentiality requirements or privacy limitations, but, at the same time, they can be more detailed than macro data, i.e. providing more breakdowns than available in macro data.[4]

2.2 Different user groups of micro data at the Deutsche Bundesbank

The overriding principle when working with micro data is compliance with the statutory secrecy and data protection requirements, and thus maintaining the confidentiality of the information submitted by the reporting agents. European and national legal provisions therefore regulate both the user group and the access channels to micro data, prescribe the required data anonymisation and oblige data providers and data recipients to maintain data confidentiality at all times.[5]

[2] See: Economic Commission for Europe of the United Nations, (2000), p.34.
[3] See: Economic Commission for Europe of the United Nations, (2000), p.31.
[4] See: Tissot, Bruno, (2015), p.1.
[5] Taken from the Annual Report of the Deutsche Bundesbank (2015).

Before different degrees of data anonymity and legal frameworks for data access are introduced, this chapter looks at the different user groups of micro data at the Bundesbank. In general, we can distinguish between three groups. Namely, these are:

1.) Internal analysts

Internal analysts are staff members of the Bundesbank. They use micro data in their day-to-day business. Most analysts work in the Directorates General (DGs) Economics, Banking and Financial Supervision, Financial Stability and Markets. Data access is often needed for ad-hoc analyses or regular reporting. Analyses mainly focus on the latest available micro data and on individual data of reporting agents.

2.) Internal researchers

Internal researchers are also staff members of the Bundesbank. They work in DGs of the Bundesbank that conduct research (e.g. Research Center, Banking and Financial Supervision, Economics or Financial Stability). In contrast to internal analysts, internal researchers use micro data in the context of research projects. As these projects often relate to Bundesbank or DG tasks, a clear distinction between “analysis” and “research” is sometimes hard to define. However, in contrast to internal analysts, data access in the context of research projects is usually needed for a much longer time span and analyses often focus on long, panel-structured time series of available micro data (often integrated data from different sources).

3.) External researchers

External researchers are researchers who are not affiliated with the Bundesbank, i.e. they are not staff members. They work at independent scientific institutions and need access to Bundesbank micro data for research projects.

Chapter 4 looks at the legal framework for data access. Chapter 6 has a more detailed focus on data access conditions for external researchers.

2.3 Degrees of micro data anonymity

In many cases, the possibility of data access is subject to the condition that the data have a certain degree of anonymity. As previously mentioned, the micro data collected at the Bundesbank are often not anonymised in their “raw” form. This is the case because data collecting divisions often need to be in contact with the reporting agents, for instance in order to discuss reported values. With some Bundesbank tasks, for example micro prudential supervision, it is also necessary to have data that is not anonymised to monitor behaviour of certain banks / MFIs. Before these micro data can be shared, they have to be anonymised to some extent. In doing so, a distinction must be drawn between the following degrees of anonymisation:

A) Absolute anonymity

Absolutely anonymised micro data are modified to such an extent that identification of the reporting agents (or any other entity requiring protection) is virtually impossible. Modification is done by coarsing (e.g. microaggregation techniques) or removing individual features.[6]

B) Factual anonymity

Individual records (micro data) are considered factually anonymised if de-anonymisation is only possible by investing a disproportionately high expenditure of time, costs and labour. Thus, factual anonymity does not rule out the possibility of de-anonymisation. Instead, the concept takes into account how costly de-anonymisation would be.

Usually, factual anonymity uses methods of information reduction or alteration to make de-anonymisation as hard as possible (aggregation, using classes instead of individual values, imputation). At the same time, attempts are made to preserve the analytical value of the data at hand. These anonymisation techniques are complemented by all technical and organisational measures taken in order to protect the data (on-site and off-site use of data, safe guest researcher workstations etc.).

Special technical and organisational measures in combination with formally anonymised data (see point C) lead to a level of data protection which is seen as equivalent to the level of factual anonymity. In such cases, data access conditions to formally anonymised data are equivalent to those of factually anonymised data. Chapter 6 provides an overview of the Five Safes framework. In that chapter, all technical and organisational measures taken at the RDSC are set into the context of one of the Five Safes categories: safe data, safe projects, safe people, safe settings and safe output.

C) Formal anonymity

Formal anonymisation involves the removal of direct identifiers. Direct identifiers are, for example, public IDs, names and addresses. Most datasets at the RDSC are formally anonymised. In combination with the protective measures taken at the RDSC (e.g. guest researcher workstations) formally anonymised data are seen as de facto factually anonymised.

[6] Analyses results which have undergone the disclosure control at the RDSC (see Chapter 6) are seen as absolutely anonymised.

Figure 2 illustrates the relationship between different degrees of anonymisation, their analytical potential and data access possibilities and takes a short look at what is laid out in the following chapter on the legal framework for data access:

- Access to original ("raw", not anonymised) micro data is very restrictive and often limited to purposes defined in the data collecting regulations.
- In some cases, data access to (de facto) factually and formally anonymised micro data may be provided to academic institutions for the purpose of academic research if the respective data can only be traced to their source with a disproportionately large amount of time, costs and labour or if special technical and organisational data protection measures are taken. It follows that, for external researchers, access to these data can only be granted on-site and that analysis results are subject to a mandatory disclosure control. The same holds for internal researchers whose research project is not needed to fulfil Bundesbank tasks.
- Absolutely anonymised data are, in principle, open for general use.

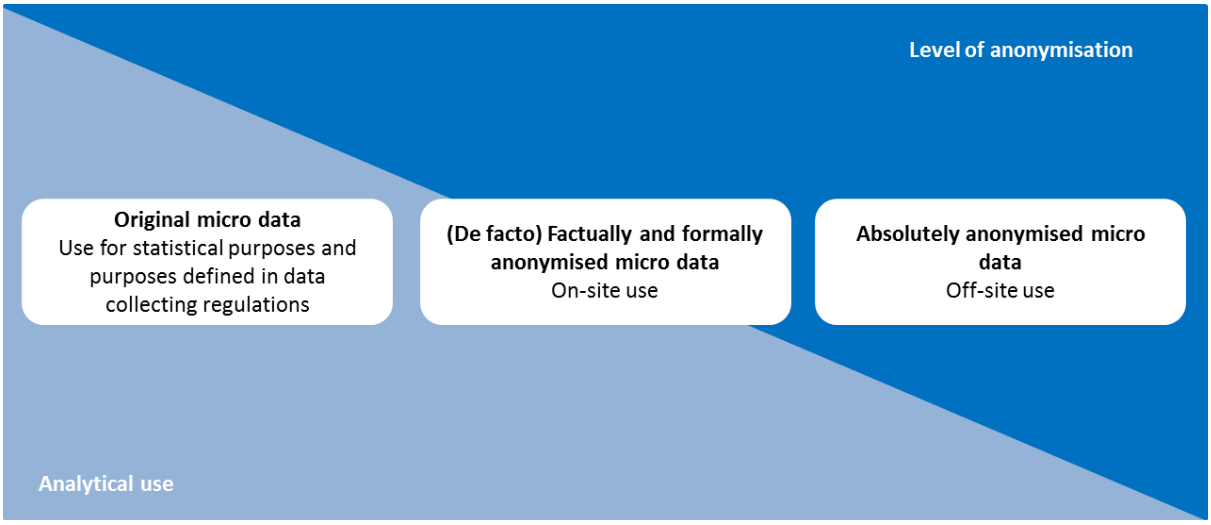
Level of anonymisation
Original micro data
Use for statistical purposes and purposes defined in data collecting regulations
(De facto) Factually and formally anonymised micro data
On-site use
Absolutely anonymised micro data
Off-site use
Analytical use

Figure 2: Analysis potential, data anonymisation and data access (own illustration following Deutsche Bundesbank, Annual Report (2015), p. 58).

3. IMIDIAS - An initiative for micro data sharing at the Deutsche Bundesbank

Most of the data collected and processed at the Bundesbank are stored and administered in separate, independent operational systems. This often results in specific isolated solutions for database infrastructure, software and hardware. In order to make micro data of these data silos available across different divisions within DGs and among DGs, the Bundesbank launched the *IMIDIAS* initiative (Acronym for: Integrated Micro Data-based Information and Analysis System).[7]

IMIDIAS provides an integrated, cross-divisional information system for analysis and research purposes. The system is micro data-based and aims at different user groups within the Bundesbank, e.g. researchers and analysts from various DGs. IMIDIAS consists of three components: the House of Microdata (HoM), the Research Data and Service Centre (RDSC) and an integrated data management process through common identifiers and data models. Internal users of IMIDIAS may be granted direct access to the HoM. External researchers can only access micro data through the RDSC.

The HoM is built into the already existing database infrastructure of DG Statistics (i.e. the database for macro data). According to a decision taken by the IMIDIAS Steering Committee, selected datasets from the above-mentioned “data silos” are transferred to the HoM. Within the HoM, the globally established ISO standard Statistical Data and Metadata Exchange (SDMX)[8] is used to classify data. This helps in structuring the data and eases comparability and linking of data from different data sources. With the concept of decentralised data responsibility and a system of staged user rights, each division retains control of its data. The HoM does not replace operational systems of the Bundesbank (the data silos remain in place), but forms an independent analysis layer with "clean copies"[9] of the process data[10].

Since access to these sensitive micro data can only be granted under a specific legal framework and data protection has to be ensured at all times, the Bundesbank has established the RDSC. It provides internal and external researchers and internal analysts with access to selected micro data of the Bundesbank. Amongst other things, the RDSC has to ensure comprehensive documentation and archiving of the data and is also responsible for advising data users on data content and analysis options. Figure 3 illustrates the data flow in IMIDIAS.

[7] For a description of IMIDIAS, see Deutsche Bundesbank, Annual Report (2015), pp. 57-59. Further information on this topic can also be found in von Kalckreuth (2014) and Bender & Staab (2015).

[8] An extensive description of the SDMX standard can be found in Stahl & Staab (2017).

[9] Clean copies are seen as final versions of a dataset which are no longer subject to changes in the data.

[10] Process data are data in stages prior to that of a clean copy.

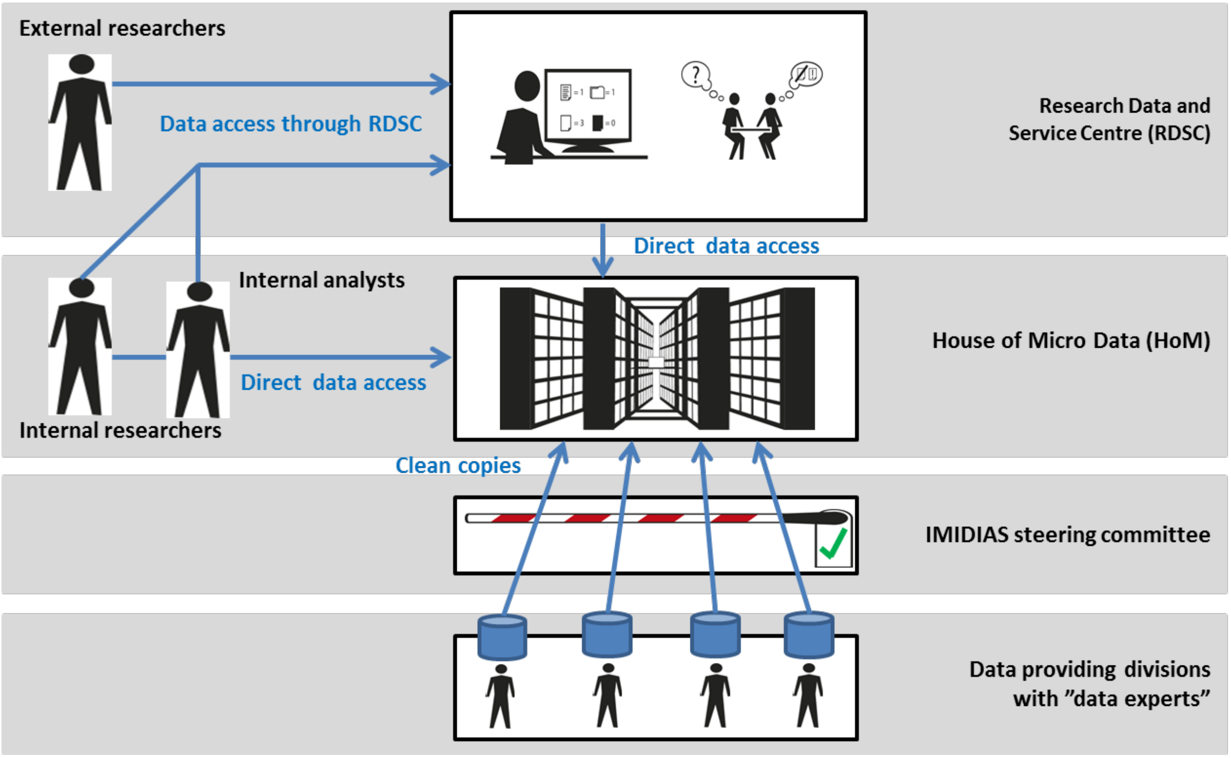
External researchers
Data access through RDSC
Research Data and Service Centre (RDSC)
Direct data access
Internal analysts
Direct data access
Internal researchers
House of Micro Data (HoM)
Clean copies
IMIDIAS steering committee
Data providing divisions with "data experts"

Figure 3: Data flow in IMIDIAS (own illustration).

4. Legal framework for micro data access at the Deutsche Bundesbank

The access to micro data at the Bundesbank and its IMIDIAS initiative is governed by binding legal standards. As each available dataset is subject to a particular legal regime, for accessing the data in question (e.g. data needed for a particular research project) the legal requirements of the specific dataset need to be met. The requirements applicable to a data request and possible limitations or restrictions arising therefrom are contingent on the individual case and in particular on the dataset in question. In the following, some non-exhaustive legal principles and fundamentals relevant for accessing IMIDIAS data are described.

As an overarching rule, access to micro data may only be granted on a need-to-know basis. This is a common principle applicable to many datasets. If the objective of a data request can be met equally without that data, access may not be granted. Similarly, where certain limitations or restrictions (e.g. anonymizations) would not hinder analyses, these would need to be implemented. Hence, in order to enable full application of the need-to-know principle, detailed explanations and justifications of the access request can be necessary.

Further requirements may arise from the applicable legal regime in a given case. Such regimes can originate in Union law, German national law or other sources (e.g. contracts and licenses). These binding regimes stipulate differing conditions and may allow non-aggregated access only for specific groups of stakeholders or certain predefined objectives. For interested parties, including independent researchers and scientists, this can result in the need for limitations and restrictions to their data access.

For instance, statistical data may either have been collected in an ESCB context or on a purely national legal basis. For data of the first category (ESCB statistical data), the relevant legal confidentiality regime is enshrined in Article 8 of Council Regulation (EC) No 2533/98 while for national statistical data collected by Deutsche Bundesbank the Bundesbank Act (Bundesbankgesetz) and section 16 of the Federal Statistics Act (Bundesstatistikgesetz) apply. This is different for banking supervisory data, the possible usage of which is outlined in section 9 of the German Banking Act (Kreditwesengesetz).

Finally, where the scope of a data request encompasses personal data or where it cannot be excluded that personal data is embraced, additional restrictions apply. Data protection legislation allows processing of personal data only under exceptional circumstances and possible legal gateway provided therein must be assessed carefully. In such cases, a high degree of anonymization may be considered.

5. Data access for internal analysts (IRMA)

In order to establish a standardized micro data access procedure for internal analysts the Bundesbank has founded a unit called “Internal Service for Microdata Analysis” (IRMA). The main goal of IRMA is to ensure that data access for internal analysts is auditable and legally compliant. IRMA is not responsible for handling data access requests from internal and external researchers. This is (still) done by the RDSC.

Before IRMA was founded, internal analysts sent their data access requests to the data providing divisions in DG Statistics. There was no overall formal and standardised process. Data access procedures were heterogeneous across different divisions in DG Statistics. With some DGs, DG Statistics had framework contracts (bilateral agreements) which regulated data sharing between these parties; with others, internal analysts had to send requests individually. Terms of Service for the usage of data from DG Statistics were agreed upon by signing a short document listing all relevant rules.

With the foundation of IRMA, the Bundesbank planned to change this situation. IRMA was given different tasks in the data access process, which should lead to more standardisation:

- Act as central point for all data access queries regarding data from DG Statistics (in the future, IRMA might also coordinate data access for other DGs).
- Advise and support data providing and data receiving divisions on the data access process.
- Develop and maintain user guides which explain the data access process.
- Check incoming data access requests for completeness and plausibility.
- Document which datasets have been approved for which purpose and by which division.
- Maintain a central data user repository.
- Professional responsibility for the respective IT applications.
- Improve consistent disclosure control regulations for research / analysis results with regard to data confidentiality on outputs / publications.

At the moment, the data access process through IRMA follows a multi-level approach:

Step 1: **General request**
Request for data access including justification for necessity of the data (one request per division).

Step 2: **Personal registration** of the data user. Agreement on the Terms of Services of data usage.

Step 3: **Specific data request** and **specific check for necessity** of the data. Data users can refer to the general request (Step 1) in their specific request.

Step 4: **Data provision** (mostly through HoM or exchange folders).

This multi-level approach replaces any existing bilateral framework agreement between DG Statistics and other DGs. All data access requests (initial requests / requests for amendment / data updates) have to run through the four steps irrespective of their data type (European / national) or the data provision channel (HoM / exchange folders).

6. Data access for internal and external researchers at the RDSC

6.1 The Five Safes framework

The Five Safes framework is a framework enabling external researchers to access detailed and unpublished data for research purposes – delivering public benefit while protecting the confidentiality of personal information. It was initially developed by Felix Ritchie at the UK Office for National Statistics and is described extensively in Desai et al. (2016).[11] Box 1 provides a brief overview of the framework.

The UK Data Service uses a security philosophy called the Five Safes framework. This is a set of principles which enable data services to provide safe access to data for research, meeting the needs of data protection, but also fulfilling the demands of open science and transparency. Five Safes has been adopted by a range of secure labs, including the Office for National Statistics.

The UK Data Service Secure Lab provides Approved Researchers with controlled access to sensitive or confidential data, enabling researchers to access and use datasets in a secure and responsible way.

The five simple protocols provide complete assurance for data owners and researchers by ensuring the following 'safes':

- Safe data: data is treated to protect any confidentiality concerns
- Safe projects: research projects are approved by data owners for the public good
- Safe people: researchers are trained and authorised to use data safely
- Safe settings: a Secure Lab environment prevents unauthorised use
- Safe outputs: screened and approved outputs that are non-disclosive

Source: https://www.ukdataservice.ac.uk/use-data/secure-lab/security-philosophy

Box 1: Overview of Five Safes framework.

The framework can be seen as a guideline for institutions that want to establish “safe” data access to their micro data for external users. At the RDSC of the Bundesbank, all of the technical and organisational measures taken in order to protect the data fit into the Five Safes concept. These measures are introduced in the following chapters.

[11] For more information see Ritchie (2008) and Desai et al. (2016).

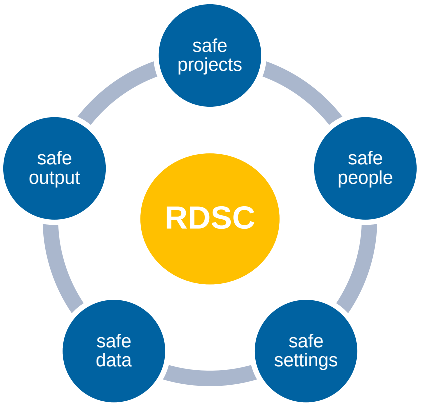

Figure 4: Five Safes at the RDSC (own illustration).

6.1.1 Safe data

Safe data mainly deals with the question of how large the potential of identification in the data is. As seen before, with Bundesbank micro data, legal frameworks only allow external researchers to access micro data that have been anonymised to some degree.

Anonymisation is done by the RDSC in collaboration with the data providing divisions. Data accessible for external researchers are either factually or absolutely anonymised. Factual anonymity is reached by giving external researchers access to formally anonymised data only in the secure working environment at the RDSC. External researchers are not given access to original data.

The legal framework for data access in combination with data anonymisation ensures data protection and thus ensures safe data at the RDSC.

6.1.2 Safe projects

The concept of safe projects mainly refers to the legal considerations surrounding the use of the data.[12] This means that data owners have to check whether data use conforms with the respective legislation. Questions that need to be answered in this context are, for example:

- Does the researcher require the data to conduct his / her project (necessity check)?
- Is the purpose of data use covered by the legislation?
- Is the researcher affiliated with an independent research institution?
- Is there any link to commercial interest in his / her research?

In order to collect the relevant information to answer these questions, interested researchers must submit an application to the RDSC prior to their first visit. The application includes a short cover letter, a CV (curriculum vitae) with a list of research activities as well as an appli-

[12] See Desai et al. (2016), p. 8.

cation form. Amongst other things, the application form asks for personal details and professional affiliation of the external researcher as well as for a project description (including motivation, planned methods, hypotheses etc.) and the data to be used.

At the Bundesbank, a two-staged approach is used when granting access rights to researchers. The data owning division has to approve the project before the RDSC can grant access rights. This way the data owner is able to monitor the extent to which their data is used.

6.1.3 Safe people

At the UK Data Service, external researchers have to pass a safe people test in order to be able to access any data.[13] The test consists of four parts:

1.) Researchers need to come from a trusted academic institution.
2.) Researchers have to demonstrate suitability for data access.
3.) Researchers need to sign a user agreement.
4.) Researchers have to undergo a one-day training course which explains legal and ethical responsibilities.

The approach used at the RDSC is quite similar. As described in the previous chapter, researchers must be affiliated with a research institution and there must not be any link to commercial interest in their research. This is checked based on the information the researcher provides in the application form and his / her CV (point 1). To demonstrate suitability, the RDSC has set a minimum educational level candidates must fulfil, i.e. a bachelor's degree (point 2).

Upon the first visit to the RDSC, researchers undergo a formal undertaking, which informs them of their legal obligations during and after the research project. The formal undertaking is conducted in person. A document listing legal obligations regarding data protection is signed by the researcher and an RDSC staff member. In addition, researchers have to sign a contract (point 3). In contrast to the UK Data Service, there is no one-day training course as yet.

6.1.4 Safe settings

Safe settings relates to the practical controls on the way that the data is accessed.[14] At the Bundesbank, external researchers can access micro data through the following channels:

[13] See https://www.ukdataservice.ac.uk/use-data/secure-lab/security-philosophy

[14] Taken from Desai et al. (2016), p. 12.

6.1.4.1 On-site access

The Bundesbank provides visiting researchers with workstations where they can access anonymised micro data on banks, securities, investment funds, enterprises and households. These workstations are (physically) located on the premises of the Bundesbank. The RDSC has 12 workstations in the TRIANON building (on the 32nd floor). On-site access is the main form of data access at the Bundesbank.

Picture 1: Workstations for external researchers at the RDSC.

In addition to the workstations at the RDSC, there are some workstations in other departments of the Bundesbank and (as of April 2019) two workstations located at the regional office of the Bundesbank in Düsseldorf. Researchers are free to choose the location of the workstations from which they want to work, but they have to inform the RDSC at least a few days in advance as permission to use the workstations is ultimately granted by the RDSC.

6.1.4.2 Technical features of a guest researcher workstation

Since the main access mode to Bundesbank data is on-site access, the RDSC must provide a “safe“ working environment for external researchers.[15] In protecting data, the main focus is to safeguard the workstations that are accessible to researchers at the RDSC. This task requires co-operation mostly with DG IT since a lot of the protection mechanisms are of a technical nature. Guest researcher workstations have a reduced functional range compared to standard Bundesbank workstations. For example, a guest researcher workstation has:

- no external interfaces (USB-ports, CD/DVD etc.)
- no internet access
- no external e-mail communication
- no printing

[15] Depending on the analysed datasets, internal researchers might also have to use a researcher workstation.

For analyses at workstations, researchers can use the following software:

- STATA (used in most cases)
- R
- MATLAB
- PYTHON

Code sharing is generally possible, but there is no formal process or platform for it. It is therefore up to the individual researcher whether he / she wants to share code with other researchers. However, the RDSC is working on a platform, where in the near future researchers will be stimulated to share their codes with others.

6.1.4.3 Off-site access

Controlled Remote Execution (CRE)

The RDSC offers controlled remote execution (CRE) for some of the micro datasets provided. Via CRE a researcher can analyse the data without being physically present at the premises of the RDSC.

On request, the RDSC sends a data structure file to the researcher, which structurally resembles the original data but contains no real values. Together with a prototype program code provided by the RDSC, this can be used to write the researcher’s own program code, which performs the desired analyses. This code is then sent to the RDSC via email and used by the staff to generate analysis results. These results are subject to statistical disclosure control. After they are checked, they are sent to the researcher by e-mail. At no point does the researcher have access to the original micro dataset.

CRE is open for researchers who have demonstrated with a current or previous research project that they have sufficient experience concerning the use of the respective micro dataset.

Scientific-Use-Files (SUF)

If data users wish to analyse micro data outside Bundesbank premises (off-site), they can request specially prepared, standardised scientific use files (SUFs). At present, the following scientific use files are available through the RDSC:

- Panel on Household Finances (PHF)
- OECD/INFE Survey of Adult Financial Literacy Competencies in Germany[16]

[16] Descriptions of data available at the RDSC can be found under: https://www.bundesbank.de/en/bundesbank/research/rdsc/research-data/research-data-617996

6.1.5 Safe Output

Together with the data providing department, researchers are responsible for ensuring that confidential data do not enter the public domain. Their published research results must remain completely anonymous. Publications must not permit any data to be traced back to individual statistical units such as banks, enterprises, individuals or households.

Consequently, the concept of safe output deals with any remaining data confidentiality risks stemming from the publication of analysis results. For example, researchers might state in a publication, that bank XY accounts for almost all of the variation in the data.[17] Such statements are a violation of privacy laws and must be prevented at all times. This is why any analysis result or publication a researcher wants to publish and use outside the Bundesbank has to pass a disclosure control by the RDSC.

In order to minimise the possibility that researchers publish individual or de-anonymised data, an internal working group at the RDSC has developed a guideline for "Disclosure control regulations for research results with regard to data confidentiality (output control)". This guideline is publicly available on the RDSC’s website[18]. Output control regulations are also part of the contract researchers sign before working with the data.

6.2 The RDSC workflow (internal vs. external researchers)

The following chapter provides an overview of the research project workflow at the RDSC. Since internal and external researchers both access data through the RDSC, Table 5 shows the differences in the workflow for these user groups. In general, the same standards apply for both user groups (especially relevant in checking applications to the same standard). However, some workflow steps may be omitted for internal researchers.

Practical experience at the RDSC has shown that internal researchers’ data access requests are much more complicated than those of external researchers (usually more datasets per project or linked datasets demanded), although their workflow process is leaner.

[17] See Desai et al. (2016), p.13.

[18] https://www.bundesbank.de/resource/blob/617956/1c0e2b50da0c5a4ef022b4f085fc8364/mL/regelungen-outputpruefung-data.pdf

Workflow step	External researcher	Internal researcher
Application for data access	required	required but with less personal information (e.g. no address needed)
Examination of application	required	required; applications are checked according to the same standards as applications from external researchers
Preparation of a guest researcher visit	required	not required as internal researchers work from their offices and workstations
Contract & formal undertaking	required	not required
Data Provision	through project folder only accessible at the secure working environment of the RDSC	mostly through project folder (sometimes through HoM); internal researchers can access the project folder from their offices and workstations
Researcher support	yes	yes
Workstation	guest researcher workstation	personal workstation in their office (a guest researcher workstation is only necessary in some cases)
Controlled Remote Access & Scientific-Use-Files	yes	not needed
Output control	required	required; can be conducted in the data using division
Publication control	required	required; can be conducted in the data using division
Archiving of project after end	yes	yes

Table 5: Research project workflow at the RDSC.

7. Conclusion

The technical report at hand shows the Bundesbanks' approach to micro data access. Until now, the Bundesbank has had very good experiences with the procedures for data access and the established system has proved its worth. Access procedures have developed well in the context of a fruitful cooperation between data providing departments and researchers. Nevertheless, the dimensions / steps / procedures presented here are not carved in stone. The Bundesbank reacts to new developments, experiences and findings and very quickly adapts the access procedure to new circumstances.

References

Bender, Stefan & Staab, Patricia (2015). “The Bundesbank’s Research Data and Service Centre (RDSC) - Gateway to treasures of micro data on the German Financial System”. IFC Bulletin No 41. https://www.bis.org/ifc/publ/ifcb41l.pdf (last accessed: 27.03.19).

Desai, Tanvi; Ritchie, Felix & Welpton, Richard, (2016). "Five Safes: designing data access for research". Bristol Business School Working Papers in Economics.

Deutsche Bundesbank, (2015). Annual Report. https://www.bundesbank.de/en/publications/reports/annual-reports/annual-report-2015-667098 (last accessed: 08.03.2019).

Economic Commission for Europe of the United Nations (UNECE), (2000). “Terminology on Statistical Metadata”, Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva.

Ritchie, Felix (2008). "Secure access to confidential microdata: four years of the Virtual Microdata Laboratory". Economic and Labour Market Statistics. 2:5: 29–34.

Stahl, Reinhold & Staab, Patricia (2017). “Die Vermessung des Datenuniversums”. Springer Vieweg.

Tissot, Bruno (2016). “Closing information gaps at the global level - what micro data can bring”, Bank for International Settlements eds., Combining micro and macro data for financial stability analysis, vol. 41.

Von Kalckreuth, Ulf (2014). “A Research Data and Service Centre (RDSC) at the Deutsche Bundesbank – a draft concept”. IFC-Bulletin No 37. https://www.bis.org/ifc/publ/ifcb37zd.pdf (last accessed: 27.03.19).