



IFC workshop on *“Combining micro and macro statistical data for financial stability analysis. Experiences, opportunities and challenges”*

Warsaw, Poland, 14-15 December 2015

## The Bundesbank’s Research Data and Service Centre (RDSC) - Gateway to treasures of micro data on the German Financial System<sup>1</sup>

Stefan Bender and Patricia Staab, Deutsche Bundesbank

---

<sup>1</sup> This paper was prepared for the meeting. The views expressed are those of the authors and do not necessarily reflect the views of the BIS or the central banks and other institutions represented at the meeting.

# The Bundesbank's Research Data and Service Centre (RDSC)

## Gateway to Treasures of Micro Data on the German Financial System

Stefan Bender, Patricia Staab  
Deutsche Bundesbank, Central Office, Wilhelm-Epstein-Straße 14

### Abstract

The Deutsche Bundesbank collects monetary, financial and external sector statistical data, comprehensive sets of indicators and seasonally adjusted business statistics. **Aggregated data** which are relevant for macroeconomic analysis are published in its macroeconomic time series databases. The Bundesbank is one of the largest data producers in Germany and its data of high quality. This applies also to its **micro data** - quality-tested administrative data covering the fields of banks, securities, enterprises and household finance.

In order to improve the sharing of micro data between Bundesbank departments, a **House of Micro Data (HoM)** is being created on the basis of the existing statistical data warehouse infrastructure. The underlying multidimensional data classification standard offers an ideal means of linking, comparing and consolidating micro data.

For external researchers, the Bundesbank also provides free of charge access to its micro data for research purposes. Due to legal requirements and in order to meet data protection requirements, individual data can be made available only under certain restrictions. Therefore, the Bundesbank has established its **Research Data and Service Centre (RDSC)** in order to provide researchers access to the Bundesbank's micro data in the context of independent scientific research projects.

The Centre grants access to most of the data during **research visits** at the Bundesbank in Frankfurt (Germany), where visiting researchers have the opportunity to view and analyse these data during research projects in a secure environment. Only **anonymized output** is leaving the RDSC.

RDSC staff members ensure that the micro data provided are **documented** in detail and **archived**. In addition, the RDSC conducts **supplementary methodological and descriptive research** based on the data sets created, and it **collaborates with researchers** within and outside the Bundesbank.

## Mandate

Since<sup>1</sup> its inception, the Bundesbank has provided broad and in-depth macroeconomic data to analysts, researchers and the general public. On the other hand, for quite some time little thought was given to providing anonymised individual-bank, firm, securities or household data – micro data.

However, detailed economic and financial stability issues, coupled with progress made in micro econometric modelling, have recently caused a spike in demand for granular data. In addition, plummeting storage and processing costs have made it easier to maintain and provide separately compiled individual-firm statistics calculated in accordance with data protection regulations and legal requirements.

Why are micro data relevant for central banks? In addition to their key task of monetary policy, central banks perform several regulatory and statistical functions. Research shows that key implications of monetary policy can hardly be uncovered using aggregate data:

- Effects of monetary policy differ across banks, firms, and households. Without taking such heterogeneity into account, the channels through which monetary policy affects prices are difficult to establish.
- Similar conclusions hold for other policy areas.
- The risk-taking effects of monetary policy – and thus financial stability implications – cannot be detected without using granular data.
- Assessing the effects of regulatory policies likewise requires granular data.

The Bundesbank – like other central banks – produces datasets which are highly valuable for policy analysis and research. So far, most of these datasets have been used to provide aggregate statistics and ad hoc analysis of specific policy issues. However, policy evaluation can make better use of them, as there is significant knowledge of data and institutional background.

Since systematic use of these data for policy analysis was up to now often constrained by time, IT-resources and legal restrictions, the Bundesbank launched a large-scale initiative called **IMIDIAS (Integrated Micro data-based Information and Analysis System)** aimed at making better use of existing data both for policy analysis as well as internal and external researchers.

<sup>1</sup> This paper can be seen as an update of von Kalkreuth (2014). There is a straight line from his more conceptual paper to the first realisation, which are presented here.

## The Bundesbank's IMIDIAS Initiative

In 2013<sup>2</sup>, the Bundesbank Statistics Department received the mandate to establish an integrated interdepartmental information system for analytical and research purposes based on micro data for various user groups (financial stability, research, monetary policy, supervision): IMIDIAS.

IMIDIAS aims to encourage cooperation with (internal and external) researchers, promote evidence-based policy-making and support policymaking processes.

In order to fulfil these goals, IMIDIAS provides several key components:

- An **integrated data management concept**, ideally based on a multidimensional data classification standard, to enable the formation of a consolidated data repository
- A data ware house infrastructure based on this concept and able to contain the data: The **House of Micro data (HoM)**
- A service unit dedicated to support internal and external data analysts and researchers: The **Research Data and Service Centre (RDSC)**

### The House of Micro data (HoM)

The House of Micro data stores **clean copies of micro data**, thus the pre-existing statistical processes will be left intact. Technically the HoM is based on the central **statistical data ware house** environment already in use by the Bundesbank's Statistics Department. The data sets are registered by a comprehensive inventory, identified by a common **potential analysis** and chosen by the **IMIDIAS steering committee** consisting of data owners and data users.

Since 2014, the IMIDIAS steering committee has been established and several actions regarding the HoM have been performed under its guidance:

- A coordinated and standardized **inventory list** of data assets at the Bundesbank, comprising the data of all relevant business units
- Several micro data / master data **pilot projects** successfully testing the capability of the recipient infrastructure
- A **potential analysis** of data sets on the inventory consisting of the main components relevance - sustainability - legal requirements - costs and leading to the identification of **twelve content projects** to form the first wave of HoM contents.
- Set up of a common framework for HoM content projects and an **IMIDIAS project office** to coordinate and control the activities

<sup>2</sup> See also Irving Fisher Committee on Central Bank Statistics (2015), p. 40-41 for a description of IMIDIAS and the House of Micro data.

Meanwhile the content projects have been launched and are at different stages of realisation:

| Status of Projects   |       |                                                                                                                                                 |
|----------------------|-------|-------------------------------------------------------------------------------------------------------------------------------------------------|
| Status               | Count | Contents                                                                                                                                        |
| Finished             | 1     | Profit and Loss Accounts                                                                                                                        |
| Implementation phase | 1     | Monthly balance sheet statistics                                                                                                                |
| Investigation phase  | 10    | COREP (two data sets), EADB, Corporate balance sheets, CSDB, Quarterly borrowers' statistics, Securities Holdings Statistics, TARGET2, TMS, TOP |

## The Research Data and Service Centre (RDSC)

The RDSC was established in 2014 as a part of the Bundesbank's Statistics Department and has since taken over several tasks formerly performed in the Bundesbank's Research Centre.

The RDSC offers access for non-commercial research to (highly sensitive) Bundesbank micro data and aims to be the single point of contact in this field for all internal and external researchers. In the long run its richest source of data will be the HoM, but the RDSC does not have to rely exclusively on this supply line. By end of November 2015, the RDSC's twelve employees situated in the centre of Frankfurt supervised over 150 active research projects, performing tasks like

- Generating (linked) micro data
- Offering advisory service on data selection and data advices (data handling, research potential, scope and validity of data)
- Providing data access (guest stays in Frankfurt, working with available software like Stata and SAS, partly MatLab and Gauss) and data protection (in particular a disclosure review service)
- Documenting data and methodological aspects of the data (available in English)

Besides the supervision of research projects the RDSC organizes conferences and workshops regarding micro data issues. In order to meet the researchers on equal footing the RDSC staff spend part of their time to work on supplementary methodological and descriptive research projects (in close cooperation with the Bank's business areas and the Research Centre).

The RDSC is part of a greater community: regarding the national side, the RDSC is part of the German data infrastructure (German Data Forum; RatSWD). At the international side, the RDSC has a lot of contacts to other Central Banks where there a similar plans to give access to microdata. The RDSC has presented the approach for data access for example at the NBER Summer Institute or the IMF Statistical Forum.

## Data Access in the RDSC

### Introduction/General<sup>3</sup>

The RDSC satisfies data protection requirements<sup>4</sup>, because it reviews incoming requests for access to micro data according to legal requirements. Applicants have to hand in an application form, expose and CV to get an approval by the RDSC. It goes without saying that access is only possible for non-commercial research.

It grants access to most of the data during research visits at the Bundesbank in Frankfurt (Germany), where visiting researchers have the opportunity to view and analyze these data during research projects in a secure environment. For this purpose the RDSC provides twelve working places for guest researchers without any external interface (internet, CD, USB) in Frankfurt.

Only anonymized output is leaving the RDSC, such as the SUF of the German PHF-study which is sent in an encrypted email.

In the near future the RDSC aims to also provide remote execution – still, relating to the complexity of the data, a guest stay in advance will always be required.

### Excursion: Methods and Applications to ensure data security

In practice, a data disseminator like the Bundesbank has to resolve the dilemma: on the one side to offer data with the largest possible information potential for research, on the other side to guaranty privacy for the units in the data.<sup>5</sup>

There are two main concepts for solving the dilemma (and of course the combination of both):

- statistical disclosure control techniques and/or
- access through a Research Data Centre (RDC)

The term **statistical disclosure control** refers to concepts and methods relating to ensure the confidentiality of micro and aggregated data that are to be published.<sup>6</sup> Statistical disclosure control refers to methodology used to design

<sup>3</sup> Some relevant paragraphs were taken out of the papers by Hochfellner, Müller and Schmucker (2014) and/or Hochfellner et al. (2012) and translated into the conditions of the Deutsche Bundesbank.

<sup>4</sup> For a description of the institutional framework, see Irving Fisher Committee on Central Bank Statistics (2015), p.39-40.

<sup>5</sup> From a legal point of view, two conflicting constitutional principles operate: (a) the personal right to informational self-determination (privacy) and (b) the academic freedom.

<sup>6</sup> An elaborate description of methods concerning statistical disclosure control can be found in Willenborg and de Waal (2001). Various international research collaborations are engaged in the development of technical solutions addressing data confidentiality (e.g., Hundepool et al., 2010). Various methodological research endeavors work on improving statistical disclosure control procedures (e.g., Drechsler & Reiter, 2009).

statistical outputs in a way that someone with access to that output can't related a known individual (or other responding unit) to an element in the output. To give a simple example, if means and percentages are presented for a large number of people, it will be impossible to infer an individuals' value from such output, even if one would know that a certain individual or unit contributed to the formation of that mean or percentage. The probability to infer information about an individual unit from statistical output can vary. If information is presented for subgroups or in multivariate tables with small cell sizes, the risk for disclosure increases. Thinking this through it becomes obvious pretty quickly that researchers are limited if they would be restricted to the use of data disseminated with these statistical disclosure techniques.

For that reason, in Germany and other countries more and more **Research Data Centres (RDCs)** had been established over the past two decades. Here qualified researchers gain access to micro-level data after they are sworn in to protect the confidentiality of the data they access and access is only given within a secure computing environments, for instance, in a RDCr (Lane, Heus, & Mulcahy, 2008). Strong input and output control is in place to ensure that published findings comply with the privacy regulations. Because some data are still too confidential to be released as micro data, or because access should be given broader and not just to those being able to visit RDCs, sometimes synthetic data are created. Here key features of the original data are preserved but noise is added to the data so that no individual record can be found exactly (Drechsler 2011). Examples for such synthetic data are the public use file for the Survey of Income and Program Participation (SIPP).

The main task of data disseminators is to safeguard anonymity of the statistical units being analyzed. This is always associated with the aggregation level of the information that has to be protected. In general, the spectrum of degrees of anonymization ranges from the original data product to aggregated statistics. Restrictions on data access depend on the anonymization strategy that is used. Aggregated statistics may be publicly available on the Internet, whereas restricted data should only be transmitted with approval and verification. The relationship of anonymity and accessibility is outlined in Figure 1. It shows the broad range between original data and aggregated statistics and between unrestricted and restricted data access.

These aspects determine the scope of action of data disseminators in Germany, but not only in Germany. Being a data disseminator of restricted data in an international network, a RDC has a somewhat smaller radius of action. In general, following rule applies: The higher the degree of anonymity, the more flexible the ways of data access that can be used. For instance, sensitive information requires strongly regulated data access. Thus, for all different kinds of data access possibilities, a RDC ensures data security by means of various procedures.

To coordinate the procedures, a RDC – and so the RDSC - implements a portfolio approach following Lane et al. (2008) and Ritchie (2011). Basically, the RDSC distinguishes between measures implemented prior to data usage, and those that take place following data usage. Prior to data usage, access is regulated by data use agreements. In addition, data are anonymized depending on the kind of data access. After data usage the results of analyses conducted using restricted data undergo disclosure review to verify compliance with data protection legislation.

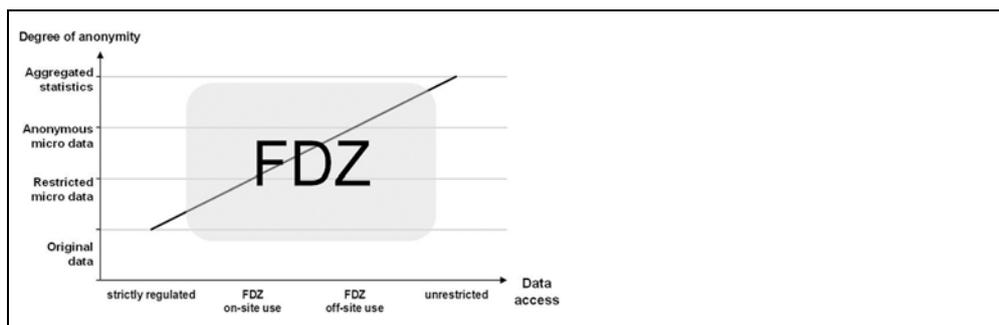


Figure 1. Degree of anonymity and data access at the RDC (FDZ: Forschungsdatenzentrum)

### Disclosure Control Prior to Data Usage

Lane et al. (2008) suggest a portfolio approach to achieve data protection while granting data access to researchers. They distinguish data protection according to four subfields, namely, technical, organizational, statistical, and legal protections. The Bundesbank's RDSC adapts their portfolio approach, which is in line with the basic principles of data disseminators worldwide, and therefore well suited to establish data protection strategies within secure computing environments.

#### Eligibility of data usage:

In most secure computing environments, access to confidential micro data is conditioned on specific regulations the data disseminator must comply with. In accordance with the German legal regulations, the use of RDSC data is linked to certain conditions. To clarify whether these conditions are met, a request for data access has to be submitted.

#### Regulations on data access:

After the data request has been approved, data use agreements are developed in which the conditions for using them are regulated. Data usage is only permitted for the specific project within the period stipulated in the agreement. The researchers who are entitled to access the data are specified; this group of people is to be kept as small as possible. All data use agreements contain bans on disclosing data to third parties, linking the data to other micro data, and identifying individuals.

### Disclosure Control after Data Usage

Aggregated data in which it is impossible to identify either individuals or establishments or banks are classified as absolutely anonymous. However, results from aggregated micro data displayed in tables are not automatically considered absolutely anonymous. If, for example, displayed cells contain only one bank or firm or person that aggregated table is not classified as anonymous. As problems like this frequently arise when analyzing restricted data, all results have to undergo disclosure review prior to releasing research output. The control of research outputs cannot be integrated entirely into a standardized and automated procedure (Gomatam, Karr, Reiter, & Sanil, 2005). Instead, the statistical disclosure control must always be tailored to the individual case at hand (Ritchie, 2011).

Prior to the publication of every research paper containing results from analyses conducted at the RDSC, the results are subject to disclosure control by the RDSC for the reidentification of individual observation units. Prior to submitting a research paper, the RDSC therefore asks every researcher to check whether it contains any micro data. If the disclosure control by the RDSC finds that it still contains micro data, the researcher will then be required to submit a corrected version of the research paper. At least five working days are needed for every submission and resubmission for disclosure control.

## A closer look at the data

The Bundesbank's micro data cover the fields of

- Banks (e. g. monthly balance sheet statistics, external position of banks, quarterly borrowers statistics, MFI interest rate statistics),
- Securities (e. g . securities holdings statistics)
- Enterprises (e. g. Micro database Direct Investment, Statistics on International Trade in Services, corporate balance sheets), and
- Household finance (e. g. Panel on Household Finances).

Several (international) research networks draw on Bundesbank micro data, e. g. the International Banking Research Network (IBRN)<sup>7</sup> or the MiDi informal network. Also, it joins in several ECB activities such as the Competitiveness Research Network (CompNet)<sup>8</sup>. The German PHF-study data it provides pertains to the European household panel. Some datasets have already been created and compiled by the Bundesbank's Research Data and Service Centre, with more to follow. These hold great potential for the analysis of financial stability. The following paragraphs name a few examples.

- the leverage ratios of individual institutions can be examined using detailed information from the banks' **Monthly balance sheet statistics (BISTA)**. The BISTA list domestic banks' assets and liabilities on the books at the end of the month along with the analytically important breakdown of the balance sheet items by type, term and the debtor's/borrower's sector. These micro data enable us to form weighted sector-level aggregates; for instance, they could be weighted by each institution's total assets in order to incorporate the variety of institutions' sizes. This helps us to take into account the heterogeneity within the banking industry.
- The loans to enterprises and households in Germany reported in the BISTA are further broken down by sector in the **Quarterly borrowers' statistics (VJKRE)**. Balance sheet statistics and borrowers statistics, in particular, are an important element of financial stability analysis. Network models can be used to identify

<sup>7</sup> <http://www.newyorkfed.org/IBRN/index.html>

<sup>8</sup> [https://www.ecb.europa.eu/home/html/researcher\\_compnet.en.html](https://www.ecb.europa.eu/home/html/researcher_compnet.en.html)

the bilateral relationships between banks or between banks and insurance corporations. On this basis, risk transmission can be observed in detail through the credit channel within the networks, for example.

- The **External position of banks (AUSTA)** records on a monthly basis the claims and liabilities of banks in Germany and their foreign branches and subsidiaries vis-à-vis foreign banks, enterprises and households, as well as general government according to recipient country. These data enable us to assess how macroprudential regulatory measures in other countries affect German banks' business with non-residents.
- The **MFI interest rate statistics (ZISTA)** measure the interest rates applied by domestic banks in Germany and the corresponding volumes of euro-denominated deposits and loans to euro-area resident households and non-financial corporations. Information is collected on month-end holdings and intra-month new business alike. One purpose for which this information is important is analysing real estate in order to identify the potential financial stability risk of lending in this sector and in a test scenario assuming a possible loan default.
- The **Corporate balance sheets (USTAN)** obtained from non-financial companies for Bundesbank refinancing operations provide some indication of the probability of company defaults, and this information can be used to analyse the probability of credit defaults and the resulting risks to the banking sector.
- Household assets and debts are a key aspect in the assessment of an economy's financial stability. For example, loan-financed property purchases in the private sector in the USA or Spain weighed on banks' balance sheets through the credit channel amidst increased defaults among these borrowers. The analysis and assessment of systemic risk to the German financial system stemming from the residential real estate market is central to the ongoing work of the German Financial Stability Committee. The Deutsche Bundesbank's **"Panel on Household Finances" (Private Haushalte und ihre Finanzen, or PHF)** survey supplies information on households' balance sheets, including their debt, pension plans, saving and income, data on employment, consumption, attitudes and expectations with regard to the economic situation, and a large number of demographic characteristics. The data enable a better understanding of issues such as saving and consumption, the distribution of wealth, insolvency risks or the characteristics of mortgage borrowers.
- The **Securities Holdings Statistics (WP-INVEST)** capture the total securities holdings of all deposit account-keeping banks domiciled in Germany at ISIN level, broken down by customer group. When a company founders or fails, this affects not only creditors but also shareholders or holders of other debt issued by that company. Micro data on securities indicate to some degree where, what volume of and in which sector securities are being held, and thus enable us to analyse contagion in securities markets. This is necessary for assessing the stability not only of the banking sector but also of the insurance sector, where securities make up the lion's share of investments.

## Conclusion

Accessing data through the Bundesbank's RDSC is the best way for using highly sensitive data. The Bundesbank hopes to (further) stimulate international research and policy analysis with these data. Accessing the data is comparatively easy and inexpensive.

Further information on the Bundesbank's Research Data and Service Centre, the anonymised datasets provided for independent scientific research and on the special access channels required for data protection purposes may be found on the website [www.bundesbank.de/fdsz](http://www.bundesbank.de/fdsz) or via email to [fdsz@bundesbank.de](mailto:fdsz@bundesbank.de)

## References

Brandt, M., Franconi, L., Guerke, C., Hundepool, A., Lucarelli, M., Mol, J., Ritchie, F., Seri, G., Welpton, R. (2010). Guidelines for the checking of output based on micro data research: Final report of ESSnet sub-group on output SDC. Eurostat. Retrieved from [http://neon.vb.cbs.nl/casc/ESSnet/guidelines\\_on\\_outputchecking.pdf](http://neon.vb.cbs.nl/casc/ESSnet/guidelines_on_outputchecking.pdf)

Drechsler, J. (2011). Synthetic datasets for statistical disclosure control \* theory and implementation. (Lecture notes in statistics, 201), New York: Springer, 138

Drechsler, J., & Reiter, J. P. (2009). Disclosure risk and data utility for partially synthetic data: An empirical study using the German IAB establishment survey. *Journal of Official Statistics*, 25, 589-603.

Gomatam, S., Karr, F., Reiter, J. P., & Sanil, A. P. (2005). Data dissemination and disclosure limitation in world without micro data: A risk-utility framework for remote access analysis servers. *Statistical Science*, 20, 163-177.

Hochfellner, D., Müller, D., Schmucker, A. (2014): Privacy in confidential administrative micro data \* implementing statistical disclosure control in a secure computing environment. In: *Journal of empirical research on human research ethics*, Vol. 9, No. 5, 8-15.

Hochfellner, D., Müller, D., Schmucker, A., & Roß, E. (2012, June). Data protection at the Research Data Centre. RDC-Methodenreport.

Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Lenz, R., Naylor, J., ... De Wolf, P.-P. (2010). Handbook on statistical disclosure control.

Irving Fisher Committee on Central Bank Statistics (2015). Data-sharing: issues and good practices, Report to BIS Governors prepared by the Task Force on Data Sharing, January 2015.

Lane, J., Heus, P., & Mulcahy, T. (2008). Data access in a cyber world: Making use of cyberinfrastructure. *Transactions on Data Privacy*, 1, 2-16.

Ritchie, F. (2011). Statistical disclosure detection and control in a research environment (WISERD DATA RESOURCES WISERD/WDR/006). Retrieved from [http://www.wiserd.ac.uk/wp-content/uploads/2011/12/WISERD\\_WDR\\_006.pdf](http://www.wiserd.ac.uk/wp-content/uploads/2011/12/WISERD_WDR_006.pdf)

Ulf von Kalckreuth, U. (2014), A Research Data and Service Centre (RDSC) at the Deutsche Bundesbank a draft concept. IFC-Bulletin No 37, Irving-Fisher Committee on Central Bank Statistics.

Willenborg, L., & de Waal, T. (2001). Elements of statistical disclosure control (Lecture Notes in Statistics, Vol. 155). New York, NY: Springer.



Irving Fisher Committee on  
Central Bank Statistics

BANK FOR INTERNATIONAL SETTLEMENTS

---

IFC workshop on *"Combining micro and macro statistical data for financial stability analysis. Experiences, opportunities and challenges"*

Warsaw, Poland, 14-15 December 2015

## The Bundesbank's Research Data and Service Centre (RDSC) - Gateway to treasures of micro data on the German Financial System<sup>1</sup>

Stefan Bender and Patricia Staab, Deutsche Bundesbank

---

<sup>1</sup> This presentation was prepared for the meeting. The views expressed are those of the authors and do not necessarily reflect the views of the BIS or the central banks and other institutions represented at the meeting.



## The Bundesbank's Research Data and Service Center

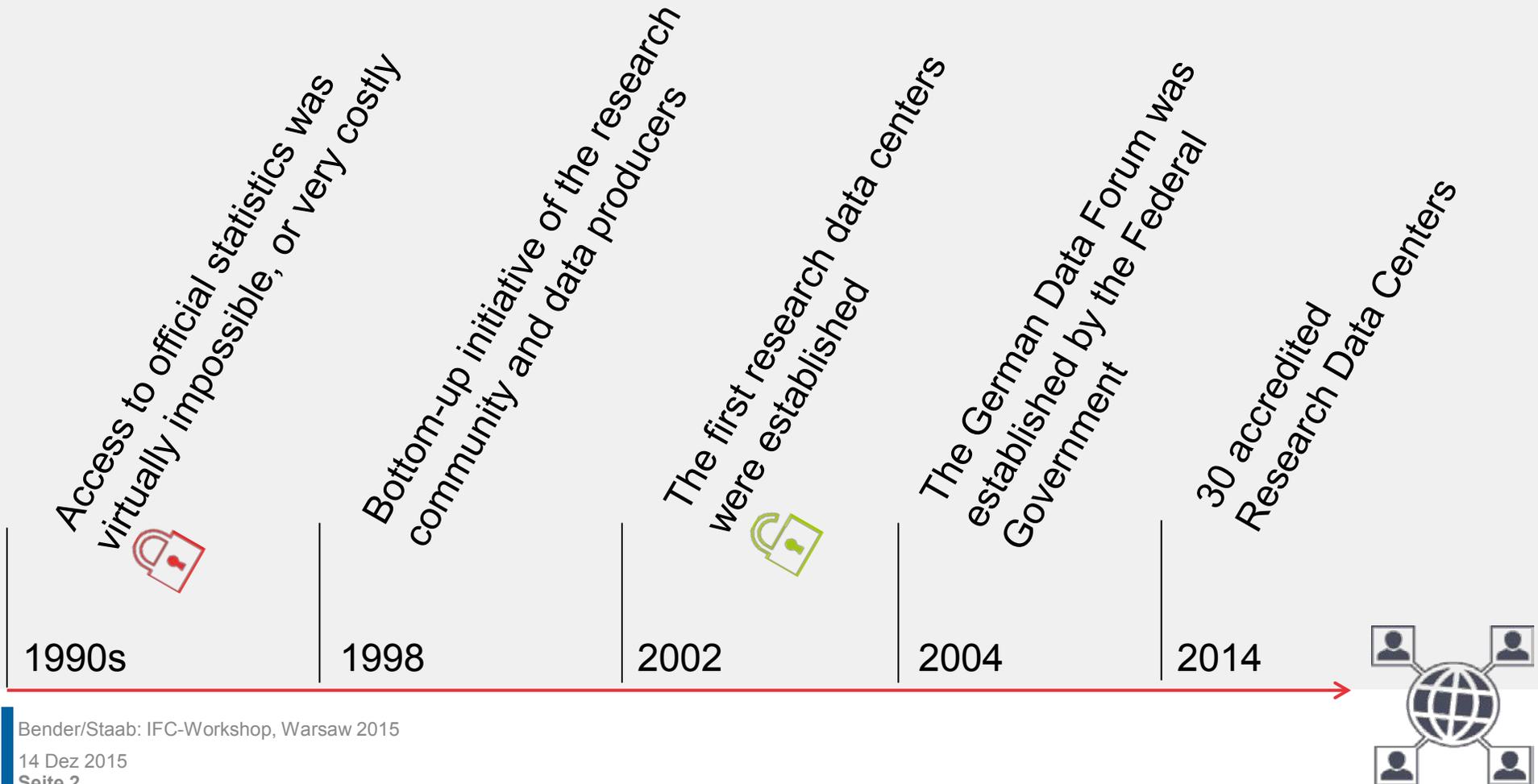
### Gateway to Treasures of Micro Data on the German Financial System

Stefan **Bender**, Patricia **Staab** (Deutsche Bundesbank)

Irving Fisher Committee Workshop on „*Combining micro and macro statistical data for financial stability analysis. Experiences, opportunities and challenges.*“

Warsaw, 14–15 December 2015

# Development of Data Access in Germany



- Relevance of micro-data for central banks
- Introduction into IMIDIAS
- Some micro-data of the Bundesbank
- Facts about the RDSC
- Conclusion

# Why are micro-data relevant for central banks?

- In addition to their key task of monetary policy, central banks perform several regulatory and statistical functions.
- Research shows that key implications of **monetary policy** can hardly be uncovered using aggregate data:
  - Effects of monetary policy differs across banks, firms, and households. Without taking such **heterogeneity** into account, the channels through which monetary policy affects prices are difficult to establish.
- **Similar conclusions hold for other policy areas.**
  - The risk-taking effects of monetary policy – and thus **financial stability** implications – cannot be detected without using granular data.
  - Assessing the effects of **regulatory policies** likewise requires granular data.

# Policy evaluation can make better use of existing datasets

- **The Bundesbank – like other central bank – produces datasets which are highly valuable for policy analysis and research.**
  - So far, most of these datasets have been used to provide aggregate statistics and ad hoc analysis of specific policy issues.
  - There is significant knowledge of data and institutional background.
- **Systematic use of these data for policy analysis is often constrained by**
  - Time
  - IT-resources
  - Legal restrictions
- **The Bundesbank has launched a large-scale initiative aimed at making better use of existing data both, for policy analysis as well as internal and external researchers.**

# Scope of the Bundesbank's Research Data and Service Center (RDSC)

- The RDSC is part of the Bundesbank internal project **Integrated MicroData-based Information and Analysis System (IMIDIAS)**
- **Goals of IMIDIAS:**
  - Encourage cooperation with (external) researchers
  - Promote evidence-based policy-making
  - Support policymaking processes
- **Key principles:**
  - Data as a public good
  - Democratic data access
  - Data protection

# Integrating Micro Data – Building a House of Micro Data

IMIDIAS is based on an Integrated MicroData Warehouse: **House of Microdata (HoM)**.

Data Sets are chosen by a bank-wide **IMIDIAS Steering Committee** (data owners and data users).

Actions:

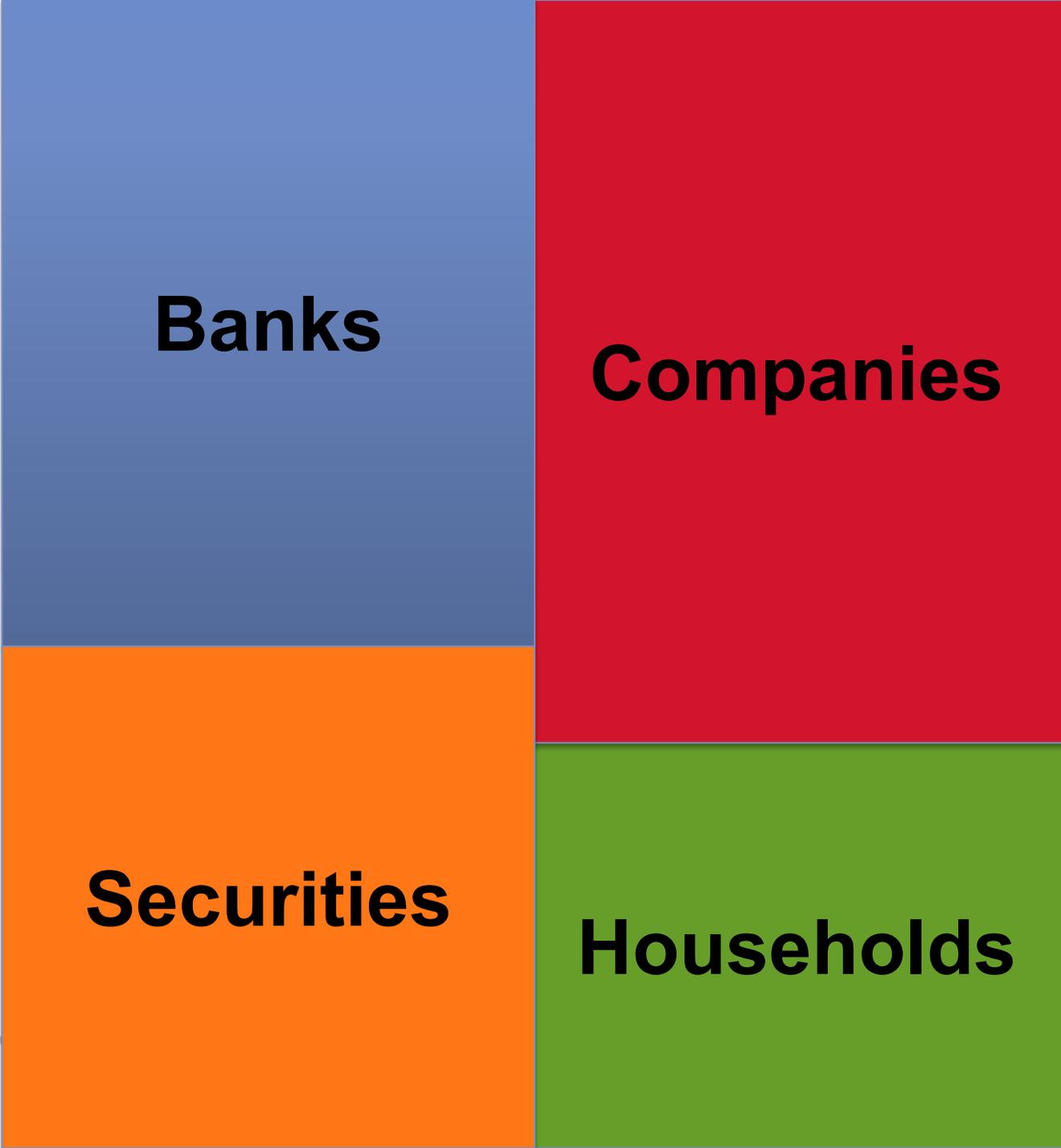
- coordinated and standardized **inventory list of data assets** at the Bundesbank
- **Pilot projects** Bank Lending Survey, Banks' Master Data
- **Potential analysis** (assessment of potential based on the inventory): Identification of **12 content projects** to create the first HoM-contents

Status of projects:

|                      |    |                                                                                                                                             |
|----------------------|----|---------------------------------------------------------------------------------------------------------------------------------------------|
| Investigation phase  | 10 | 2 COREP data sets, EADB, Securities Holdings Statistics, CSDB, Corporate balance sheets, Quarterly borrowers' statistics, TARGET2, TOP, TMS |
| Implementation phase | 1  | Monthly balance sheet statistics                                                                                                            |
| Finished             | 1  | Profit and Loss Accounts                                                                                                                    |

## Integrating Micro Data – How to choose the relevant data

- **Relevance:** Benefit for the areas of monetary policy, financial stability, payments, banking supervision, research.
- **Sustainability:** Structural stability and maturity of the data, lasting importance for users.
- **Legal requirements:** conditions for data use (internal analysis, for internal and external research).
- **Costs:** Effort necessary to put the data to use.



**Banks**

**Companies**

**Securities**

**Households**

### **Monthly balance sheet statistics (BISTA)**

- Domestic banks' assets and liabilities (monthly)

### **External positions of banks**

- Assets and liabilities of banks in Germany and their foreign branches and subsidiaries vis-à-vis non-residents.
- Input for the BIS Banking Statistics

### **Banks' loans to enterprises and households in Germany (BISTA)**

- Break down by sector, quarterly borrowers statistics

### **The MFI interest rate statistics (MIR)**

- Sample of around 240 institutions.
- Interest rates applied by domestic banks (MFIs) and the corresponding volumes for euro-denominated lending and deposit business with households and non-financial corporations in the Euro Area.

### **Annual statistics on foreign direct investment stocks (MIDI)**

- Since 1976, in accordance with the provisions of the Foreign Trade and Payments Regulation
- Affiliate-level panel data available since 1996

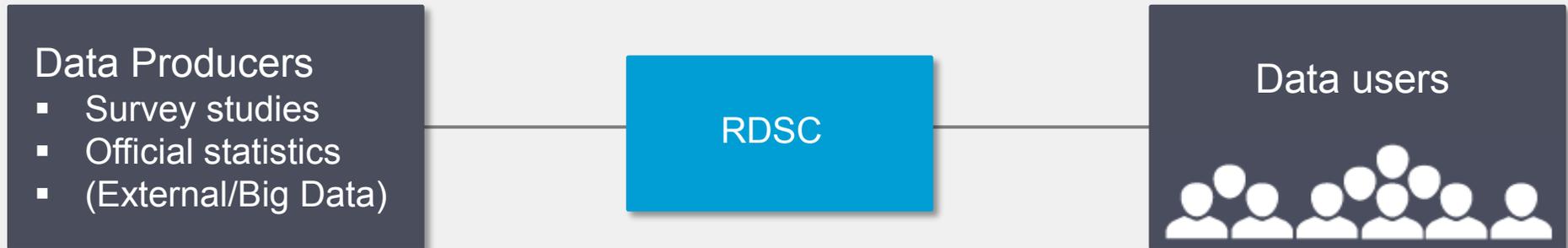
### **Germany's Statistics on International Trade in Services (SITS)**

- Part of the Balance of Payments (**BoP**) Statistics
- Information on international service transactions carried out by German residents that exceed €12,500.

### **Financial statements of non-financial companies (Ustan)**

- Originates from Bundesbank's refinancing activities
- Information on earnings and financing (1987 to 2011).

# Data Access in the RDSC



- RDSC mediates between data producers and external users.
- RDSC controls for compliance with data protection regulations.

## Tasks of the RDSC

RDSC offers access for non-commercial research to the (highly sensitive) micro data.

- Generates micro data (linking data).
- Provides data access and data protection.
- Offers advisory service on data selection and data access (handling, potential, scope and validity of data).
- Documents data and methodological aspects of data.

## Factsheet on the RDSC of the Bundesbank

- The RDSC has started in 2014 as part of the Statistics Department of the Bundesbank.
- Over 150 active projects.
- 12 employees (2016: 14).
- 12 working places for guest researchers.

# Modes of Data Access

| Off-Site Access                                                                   | On-Site Access                                                                      |
|-----------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
|  |  |
| Email, encrypted<br>(Scientific Use File)                                         | Remote Execution<br>(near future)                                                   |
| Factually anonymous                                                               | Weakly anonymous (= confidential)                                                   |

# Location of the RDSC

- 20<sup>th</sup> floor of the Trianon-Tower in Frankfurt.
- Near Frankfurt central station.



Working places for guest researchers



Trianon-Tower

## Conclusion I

- Policy advice and research is about answering questions.
- There is a strong need for granular data.
- Access to data is needed.
- Solutions are in place to fulfill privacy issues (RDSC).

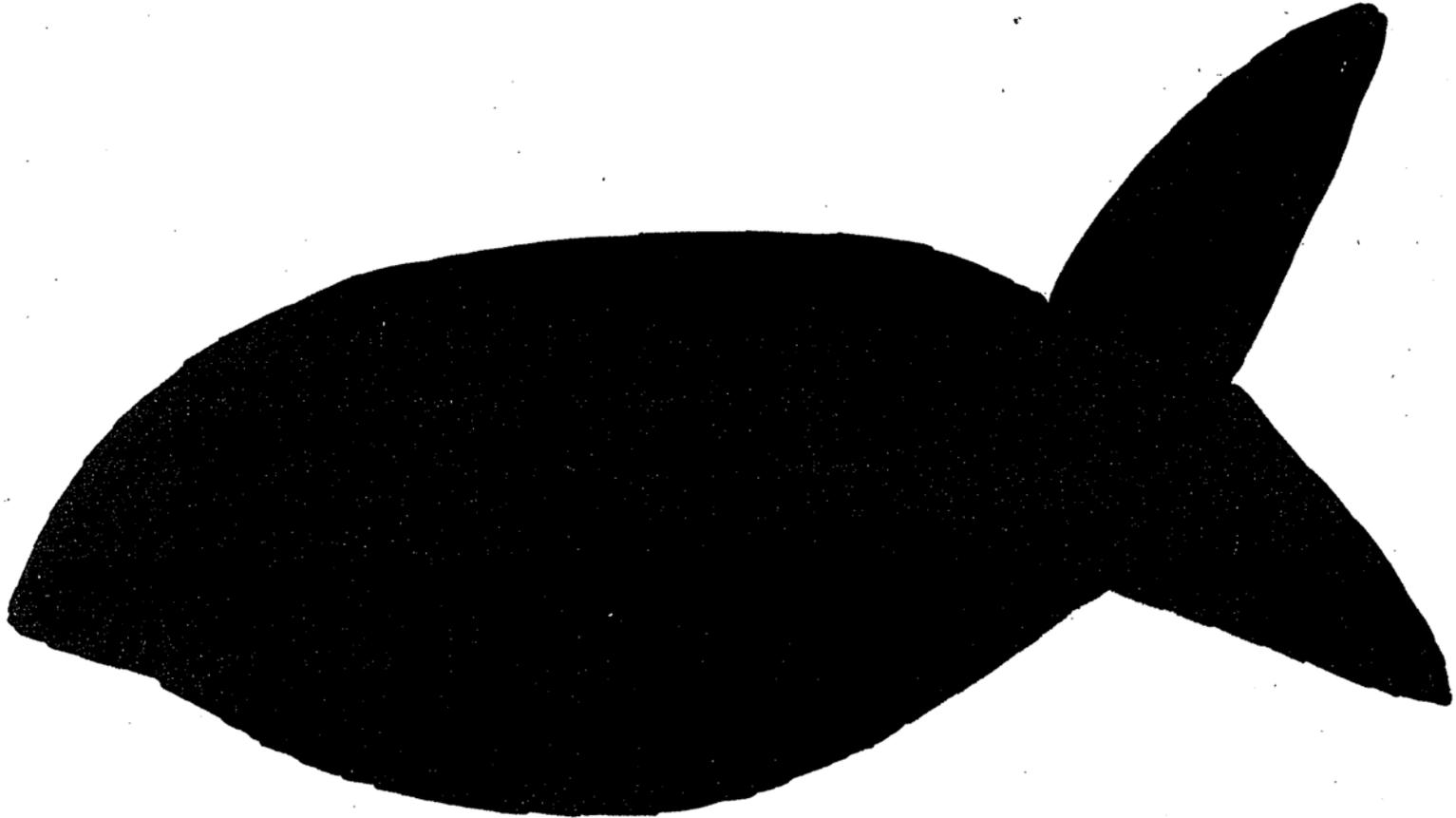
## Conclusion II

- Accessing data through a RDSC/RDC is the best way for using highly sensitive data like Bundesbank's data.
- The Bundesbank hopes to (further) stimulate international research and policy analysis with these data.
- Accessing the data is comparatively easy and inexpensive.

➤ Come, visit us in Frankfurt!

- **Website:** [www.bundesbank.de/fdsz](http://www.bundesbank.de/fdsz)
- **Contact:** [fdsz@bundesbank.de](mailto:fdsz@bundesbank.de)

# Sometimes



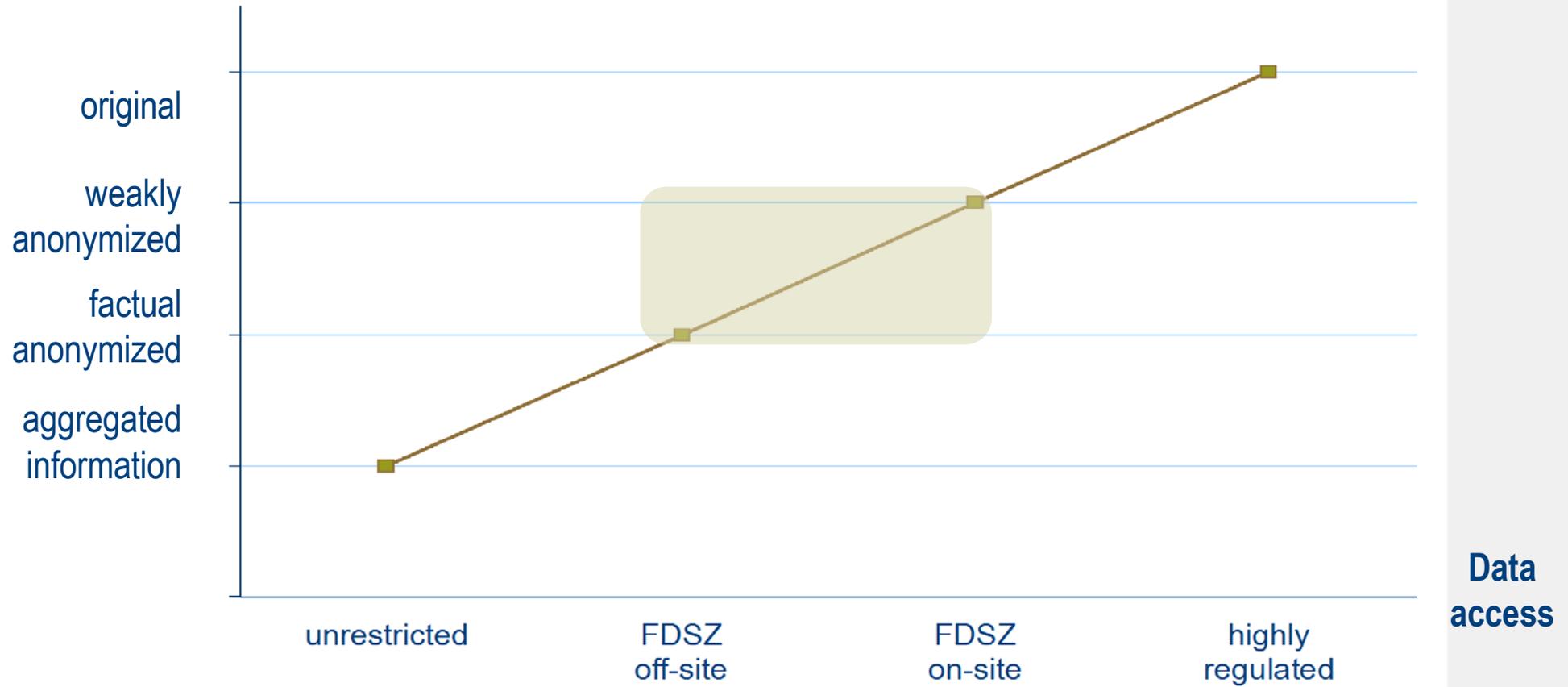
# Backup

The RDSC offers access for **non-commercial** research to (highly sensitive) micro data of the Bundesbank:

- Generate (linked) micro data
- Offer advisory service on data selection and data access (data handling, research potential, scope and validity of data)
- Provide data access and data protection
- Document data and methodological aspects of the data
- Work on own research projects (in close cooperation with the Bank's business areas and the **Research Centre**)
- Organize conferences and workshops.

# Modes of Data Access

## Potential



## Data access in Detail

- **Application form, expose and CV** are needed to get an approval by the RDSC. Access for non-commercial research.
- Micro data are only available via a guest stay (on-site).  
**Exception: SUF of the German PHF-study** (data will be send in an encrypted email).
- **12 working places for guest researchers** without any external interface (internet, CD, USB).
- Near future: Remote execution. Because of the complex data a guest stay in advance is required.

## Other services provided by the FDSZ

- Software available: Stata, SAS
- Additional software: MatLab, Gauss
- Disclosure review service
- Documentation in English
- Advisory service and assistance