Giving the International Scientific Community Access to German Labor Market Data: A Success Story

Dana Müller and Joachim Möller

1 Introduction

In Germany, as in other countries, social security data offer a great opportunity for producing cutting-edge empirical analyses. They are the basis for answering relevant research questions as well as for evaluation studies and evidence-based policy-making. Data resources are especially valuable if they are linked to establishment and individual survey data.

As the research unit of the Federal Employment Agency, the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung (IAB)) is responsible for extracting data from administrative processes to produce micro-datasets that can be used for empirical research on a wide range of labor market topics. In the past, the data were generally kept within the organization, which not only led to a drastic underutilization of the data resources but also limited collaboration projects with national and international academic scholars. There were only rare examples of knowledge spillovers from the international research community to the Institute’s research projects. With some major exceptions, researchers at the Institute faced difficulties in keeping pace with the enormous evolution of (micro-) econometric methods. As a consequence, data analysis was mainly descriptive, and publication in refereed international journals was the exception rather than the rule.
With the growing realization that a closed strategy hinders scientific progress, the strategy was already being softened in the 1990s. In addition, the scientific community showed growing demand for exploiting the valuable data resources to answer research questions. This outside pressure favored the process of opening; however, development took some time. There was no standardized and institutionalized way for researchers outside the IAB to gain access to the data until 2004 (Kohlmann 2005).

There were two important impulses to improve data access for the scientific community. The first was the labor market reforms implemented between 2003 and 2005. An element of these reforms was to strengthen scientific evaluation of active labor market instruments to increase their efficiency. The second was the recommendation of the German Commission on Improving the Information Infrastructure Between Science and Statistics (Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik) to establish a research data center at each public producer of microdata in Germany. The Federal Employment Agency (Bundesagentur für Arbeit (BA)) followed this recommendation and established a research data center within the IAB in spring 2004. This was facilitated by the Federal Ministry of Education and Research, which funded the initial process for 3 years. After an evaluation by the German Data Forum (Rat für Sozial- und Wirtschaftsdaten) in 2006, a research data center was established as a permanent department of the IAB (see Solga and Wagner 2007). Today, the Research Data Centre of the Federal Employment Agency at the IAB (RDC-IAB) is one of 31 such research data centers that have been established (see http://www.ratswd.de/en/data-infrastructure/rdc).

The establishment of the RDC-IAB and the reorganization of the institute under the directorship of Jutta Allmendinger were the starting signal for numerous collaborations with external scholars. IAB researchers profited especially from joint projects with international partners. Some of these projects led to publications in top-ranked international journals (see Dustmann et al. 2009; Card et al. 2013, among others). Active labor market policies were evaluated using the latest empirical methods (see, for instance, Wolff and Stephan 2013). In several cases, the labor market research based on the RDC-IAB data led to a new design of active labor market policy. A recent example is an evaluation study on the compulsory integration agreement between the jobseeker and the caseworker. Using a randomized field experiment and following the labor market biographies of the persons included in the experiment, IAB was able to show that for some groups of unemployed, the compulsory regulation is counterproductive and should be replaced by a more flexible handling of the instrument (van den Berg et al. 2016).

Another important field is to monitor and evaluate the effects of new labor market regulations or institutions. One example is the minimum wage that was first implemented in the German construction industry in 1997 and later extended to other sectors (König and Möller 2009; Möller 2012; Aretz et al. 2013). The various effects of the general statutory minimum wage that was implemented on 1 January 2015 are currently being analyzed in several projects based on RDC-IAB data. In general, labor administration and policy-makers have been profiting from
better insight into labor market structures and are now able to optimize labor market processes and instruments.

Through data access points that adhere to the highest standards of data security and confidence, the RDC-IAB provides researchers with access to its data resources not only in Germany but also in the United Kingdom (UK) and the United States (US). The number of users is steadily increasing. In 2016, almost one-third of all data use agreements were from a non-German facility. In the future, the RDC-IAB will expand the possibilities of data access even further. Mutual access to microdata of different European countries or linkage of different datasets also requires new technical solutions.

The aim of this chapter is to provide an overview over the activities of the RDC-IAB and developments planned for the future. It begins with a description of the core data and the modes of data access. It then describes how demand for the data evolved over time. Further infrastructural developments and research activities are described in Sects. 5 and 6. Section 7 concludes.

2 The Research Data Centre at the IAB and Its Data Resources

The RDC-IAB is primarily a service-oriented department but also conducts its own research projects and acquires grants from various foundations. It provides access to high-quality microdata for researchers in Germany and abroad, in compliance with German data protection legislation. To accomplish this aim, the RDC-IAB performs the following tasks:

- It develops specific data products with high research potential for labor market studies; this includes the necessary preparation and harmonization of the raw data as well as regular updates.
- It compiles detailed documentation of the data products and considers technical aspects of the data as well as statistical properties. It provides tools to facilitate analyses of microdata and offers individual counseling.
- To prevent re-identification of personal information, the RDC-IAB develops and applies anonymization strategies.
- It develops standardized ways for (inter)national researchers to access data.
- It promotes data products and data access through active participation in (inter)national workshops, conferences, and seminars.
- It conducts its own research on and with the available data products to improve their quality and to assess their research potential and ability to provide competent individual counseling for external researchers.

The RDC-IAB shares its activities on data access and the development of new data products, metadata, and research with an international network of research data centers, data providers, and scientific institutions.
Thirteen years after its foundation, the RDC-IAB is considered the most important supplier of German labor market microdata. Currently, 16 data products on individuals, households, and establishments are available to the scientific community. The data originate from administrative data from the social security system’s notification process and internal processes of the BA and from surveys conducted by the IAB. The RDC-IAB enlarges the research potential by linking existing data with other administrative data or surveys. In 2011, the Record Linkage Center was founded at the IAB, a joint project with the University of Duisburg-Essen that was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft). The methods developed in this context facilitate the linkages of microdatasets without a unique identifier.

The RDC-IAB offers labor market microdata on individuals, households, and establishments. These datasets are generated from three different sources: (i) register data from the social security system’s notification process, (ii) data from internal procedures of the BA, and (iii) survey data (see Fig. 1 for an overview).

The legal basis for social security data collection is provided by the German Data and Transmission Act (Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung) and the Social Act (Social Code Book IV). As part of the social security notification procedure, all employers are required to report several items and characteristics of their employees. In principle, two kinds of information are stored. The first is information that is collected for statistical purposes, and the second is information that is collected to compute the amount of social security contributions and the resulting claims. The administrative data from the internal procedures of the BA are the result of the agency’s fulfillment of tasks in accordance with the Social Code Books II and III. These are the adminis-
tration of the compulsory unemployment insurance, calculations of unemployment benefits and the corresponding entitlement periods, consultation sessions with the unemployed, placement offers, and active labor market measures. The collection of these administrative data began in 1975. The IAB generates historical data from these records and combines them into a comprehensive unique dataset, i.e., the Integrated Employment Biographies (IEB). Not all variables are available for the entire observation period. Due to changes in statutory regulations, administrative data sources start at different points in time. The RDC-IAB updates its data products regularly and offers different samples of these rich administrative data sources for research purposes.

As well as processing the administrative microdata, IAB conducts various surveys. In addition, the RDC-IAB exploits the opportunity to link these surveys to administrative data. According to German data protection rules, this is allowed if the respondents consent to the linkage (Heining 2010).

Currently, the RDC-IAB provides 16 datasets (see Table 1). A short description of selected examples for establishment, individual, and household data is given in the paragraphs below. More detailed information is available on the website of the RDC-IAB (see http://fdz.iab.de/en.aspx).

The IAB Establishment Panel (IABB) is an annual representative survey of approximately 16,000 establishments in Germany (Bellmann 2002; Fischer et al. 2009; Ellguth et al. 2014). The survey started in 1993 for West Germany and has covered East Germany since 1996. It includes only establishments with at least one employee covered by social security on 30 June of the previous year. The IABB contains various topics, such as the development of total employment, business policy, investments, export, innovations, personnel structure, apprenticeship and vocational training, recruitments and dismissals, wages, working hours, training programs, and alternating annual topics.

The IAB Establishment History Panel (BHP) is a yearly cross-sectional dataset on all establishments in Germany with at least one employee eligible for social security contributions (Spengler 2008; Eberle and Schmucker 2017). The dataset is a 50% random sample drawn from establishment identification numbers and gives information for the reference date (30 June) of each year. The panel starts in 1975 for West Germany and in 1992 for East Germany and includes between 640,000 and 1.5 million establishments per year. The BHP contains information on workforce composition such as gender, age, nationality, occupational status and qualification as well as branch of industry and the location of the establishment. Furthermore, there is information on worker in- and out-flows and indicators of establishment entries and exits (Hethey-Maier and Schmieder 2013).

The German Management and Organizational Practices Survey (GMOPS) is a novel establishment dataset provided at the RDC-IAB since September 2016 (Broszeit and Laible 2016). GMOPS, funded by the Leibniz Association, belongs to Management Practices, Organizational Behavior, and Firm Performance in Germany, a collaboration project that was jointly carried out by the IAB, the Kiel Institute for the World Economy (IfW), and the Institute for Applied Social Sciences (infas). The survey is based on the US Census Bureau’s “Management
Table 1 Data products of the Research Data Centre at the IAB

| Dataset                                                                 | Administrative Data | Survey       | Linked Data |
|-------------------------------------------------------------------------|---------------------|--------------|-------------|
| **Establishment data**                                                   |                     |              |             |
| IAB Establishment Panel                                                 |                     | ✓            |             |
| IAB Establishment History Panel (BHP)                                   | ✓                   |              |             |
| German Job Vacancy Survey of the IAB                                     | ✓                   |              |             |
| German Management and Organizational Practices (GMOP) Survey             | ✓                   |              |             |
| **Individual data/household data**                                      |                     |              |             |
| Sample of Integrated Labor Market Biographies (SIAB)                    | ✓                   |              |             |
| Panel Study’ Labor Market and Social Security (PASS)                    | ✓                   | ✓            |             |
| Working and Learning in a Changing World                                 |                     | ✓            |             |
| Biographical Data of Social Insurance Agencies in Germany (BASID)        | ✓                   |              | ✓           |
| IAB-SOEP Migration Sample (IAB-SOEP MIG)                                | ✓                   | ✓            |             |
| Employee survey’ bonus payments, wages increases, and fairness (BLoG)   |                     | ✓            |             |
| IZA/IAB Administrative Evaluation Dataset                                | ✓                   |              | ✓           |
| lidA—*leben in der Arbeit*. German cohort study on work, age, and health |                     | ✓            |             |
| S-MGA—The Study on Mental Health at Work                                 |                     | ✓            |             |
| **Integrated establishment and individual data**                        |                     |              |             |
| Linked Employer–Employee Data from the IAB (LIAB)                       | ✓                   | ✓            |             |
| Linked Personnel Panel (LPP)                                            | ✓                   | ✓            |             |
| Panel “WeLL”—Employee Survey for the Project “Further Training as a Part of Lifelong Learning” | ✓                   | ✓            |             |

and Organizational Practices Survey” (MOPS) from 2010. Large parts of the questionnaire were translated into German, and additional information has been added, for example, on work–family balance and health promotion and on sales, export, and innovation. The survey was conducted once, in the period 2014–2015. The information in the data relates to the years 2008 and 2013 and covers 1927 establishments.

The “Sample of Integrated Labor Market Biographies” (SIAB) is a 2% random sample from the Integrated Employment Biographies (Dorner et al. 2010). The employment biographies cover the period from 1975 until 2014 for West Germany and from 1992 until 2014 for East Germany. The microdata include more than 1.7 million individuals in total and cover day-exact information on sociodemographic characteristics, employment, benefit receipts and job searches, and location and establishment.
The “Panel Study Labor Market and Social Security” (PASS) is an annual household survey in the field of labor market, welfare state, and poverty research in Germany (Trappmann et al. 2013). The survey consists of two random samples. The first sample includes households and individuals receiving means-tested social assistance (the so-called Unemployment Benefit II), and the second includes any other households of German residents. The field phase of the first wave ran from December 2006 to July 2007. Both random samples are continued over time. To guarantee representativeness in each wave for Unemployment Benefit II recipients, refreshment samples of households that claimed Unemployment Benefit II for the first time were drawn for the following waves. The survey includes a personal interview with the head of household and, subsequently, personal interviews with all members of the household aged 15 or older. Persons aged 65 or older are interviewed with a reduced questionnaire. The last wave of 2015 includes approximately 13,300 persons in nearly 9000 households. More than 11,700 of these persons and more than 7800 of these households have been interviewed multiple times.

The “Linked Employer–Employee Data” from the IAB (LIAB) combines the IAB Establishment Panel with data for employees from the Integrated Employment Biographies (Heining et al. 2014). The LIAB is useful for the simultaneous analysis of supply and demand on the German labor market. There are two different versions of the LIAB. The LIAB cross-sectional model contains all waves of the IAB Establishment Panel and linked information of all employees on 30 June of a given year. The updated LIAB longitudinal model is a sample of establishments repeatedly interviewed between 2000 and 2011 and is linked to all employees who worked at least 1 day in one of the establishments included. The employment biographies begin in 1992 and continue until 2014. Additional generated variables comprise the employment and unemployment experience before 1992.

The “Linked Personnel Panel” (LPP) is another novel-linked employer–employee dataset on human resource work, corporate culture, and management instruments in German establishments (Bellmann et al. 2015). It evolved from the “Quality of Work and Economic Success” project, a collaboration between the IAB, the University of Cologne, and the Centre for European Economic Research (ZEW). It is funded by the IAB and the Federal Ministry of Labor and Social Affairs (Bundesministerium für Arbeit und Soziales (BMAS)). The project is designed to include three survey waves of employers and their employees, at 2-year intervals. The current data product contains the first two waves. In the first wave (2012/2013), 1219 establishments and more than 7500 of their employees were interviewed. The second wave (2014/2015) contains information for 771 establishments and approximately 7280 employees. The LPP Employer Survey is directly attached to the IAB Establishment Panel; therefore, all information of the IAB Establishment Panel can be included.

For each data product, the RDC-IAB provides detailed documentation in German and English. There are two publication series. The FDZ Datenreport series contains documentation of the data, changes to previous versions and information on data preparation, as well as methodological aspects on data handling. Additional information on frequencies, labels, or working tools is available on the website.
the RDC. Currently, the RDC-IAB is working on transferring from PDF format documentation to a web application for all data documentation using the DDI standard.1 The *FDZ Methodenreport* series addresses methodological aspects and problems. It may be used as a publication outlet by any author working with BA or IAB data.

## 3 Data Access

The legal basis for data access is found in §75 of the German Social Code Book X and §282 (7) of the German Social Code Book III. The need for data protection determines the ways in which data can be accessed. In general, this means that the more detailed the data are, the more restricted the access to the data is. The RDC-IAB offers four kinds of data access for the scientific community:

1. **Campus files** are fully anonymized and useful only for teaching. Users need to register and agree to terms of use before the campus file can be downloaded.

2. **Scientific use files** are de facto anonymized microdata that are submitted to scientific institutions in Germany and EU Member States within the scope of §282(7) of the German Social Act III. The information has been reduced for data confidentiality reasons to the extent that re-identification of personal information would be possible only with a disproportionate amount of time, expense, and effort (Hochfellner et al. 2014). Scientific use files are offered to researchers for research projects in the field of labor market research but not for teaching or for commercial research interests. Data security must be guaranteed by the scientific institution applying for the data.

3. The anonymization of the scientific use file restricts the research potential; therefore, the RDC-IAB offers *weakly anonymized data* (i.e., de-identified microdata) with more detailed information. Access is possible only via *on-site use* within the scope of §75 of Social Code Book X. The RDC-IAB provides separate workplaces within a secure computing environment in Nuremberg and at various locations in Germany, the USA, and the UK (Bender and Heining 2011). Within the secure computing environment, researchers have direct access to weakly anonymized data; however, they can obtain the output of their programs only after disclosure reviews by RDC staff (for details, see Hochfellner et al. 2014). On-site use is limited to research projects in the field of social benefits or labor market research.

4. **Remote execution** means that researchers prepare their programs with artificial data and upload the programs in the Job Submission Application (JoSuA), which is described in more detail in Sect. 5. Researchers never view the original data. They receive their results after a disclosure review by RDC staff. Remote execution is also possible after on-site use of the data.

---

1DDI stands for the Data Documentation Initiative. See [https://www.ddialliance.org/](https://www.ddialliance.org/).
The use of scientific use files, remote execution, and on-site access must fulfill certain requirements in accordance with the legal regulations (for more details, see Hochfellner et al. 2014). Therefore, the RDC-IAB offers standardized request forms for all data access to clarify whether or not the research purpose complies with the legal requirements. Final permission for on-site use is granted by the BMAS. After a request has been approved, a contract of data use for a specific project within a specific period is concluded between the researcher’s institution and the RDC-IAB. The contract specifies the data protection rules and severe sanctions in the event that these rules are violated.

Note that some of the datasets listed in Table 1 are available only for on-site use. Among others, this applies for all linked datasets.

4 Development of the Demand for Data Products

The RDC’s data products enjoy immense popularity in Germany and abroad. Figure 2 shows the number of users and the numbers of projects for each year since 2005. Generally, more than one researcher works on one project, and the duration of a project usually exceeds 1 year. The number of users has increased consistently over time. In 2015, for instance, the RDC-IAB reached just over 1000 users, who work, or were working, on 514 projects. In 2016, the number of users and projects was higher still.

![Fig. 2](image_url) Development of the number of data product users and number of projects, RDC-IAB, 2005–2016
Most of the users of RDC-IAB data products work at a German research institute or university, but a growing number of data requests are from international scholars. The noticeable increase in the number of users was made possible through the project “The Research-Data-Centre in Research-Data-Centre Approach: A First Step Towards Decentralised International Data Sharing” (RDC-in-RDC),\(^2\) funded from 2011 to 2013 by the Ministry of Education and Research (Bundesministerium für Bildung und Forschung (BMBF)), follow-up funded by the National Science Foundation under program SES-1326365, and financially supported by the project “Data Without Boundaries” within the Seventh Framework Programme of the European Union (Heining and Bender 2012). Before this project began, the only location with access to weakly anonymized data was the RDC-IAB in Nuremberg. Capacity was limited to five workstations. Within the project, data access was established at various locations in research data centers or institutions that offer a data protection infrastructure similar to that of the RDC-IAB. In each partner organization, there is a secure guest room, and researchers are provided data access via a secure internet connection to the RDC-IAB in Nuremberg. In principle, there are no differences between the international and German RDC-in-RDC IAB approaches in either technical implementation or the application process. However, the difference in the legal framework must be considered. This legal framework requires that only (de facto) anonymized data can be accessed from abroad. Therefore, RDC staff construct (de facto) anonymized datasets for approved projects. Figure 3 shows,

![Fig. 3](image-url)  
**Fig. 3** Contractual partners of the IAB Research Data Center by country, 2012–2016

\(^2\)This project was initiated by Stefan Bender, former head of the RDC.
Table 2  Number of projects by datasets and year, 2012–2016

| Dataset | 2012 | 2013 | 2014 | 2015 | 2016 |
|---------|------|------|------|------|------|
| SIAB    | 69   | 107  | 154  | 203  | 231  |
| LIAB    | 77   | 85   | 101  | 111  | 132  |
| IABB    | 90   | 74   | 71   | 76   | 75   |
| PASS    | 34   | 39   | 47   | 49   | 54   |
| BHP     | 28   | 30   | 37   | 41   | 47   |
| ALWA    | 26   | 29   | 28   | 24   | 23   |
| BASiD   | 7    | 11   | 14   | 21   | 23   |
| All others | 100 | 81   | 71   | 79   | 78   |
| Total   | 431  | 456  | 523  | 604  | 663  |

Note: For explanations of dataset acronyms, see text

for example, that 31% of all research projects in 2016 were from a non-German facility.

The demand for several RDC-IAB data products depends on the user’s research purposes. Table 2 shows a list of the seven most commonly requested datasets within the past 5 years. The number of projects in Table 2 differs from the number of projects in Fig. 1 because the number of data products within a project is not restricted to one data product. The SIAB has been the most requested dataset since 2013. The SIAB is available both as a scientific use file and for on-site use. It includes individual information covering the period since 1975 and is enriched by several characteristics of the employer. The SIAB, therefore, suits a wide range of research purposes in the fields of labor market research and social research. It comes as no surprise that the Linked Employer–Employee dataset LIAB ranks second. Here, valuable and reliable information from the administrative data is combined with comprehensive information from the establishment panel. These rich and innovative data hold enormous research potential.

The figures and tables above show the high demand for RDC-IAB data products. The importance for the scientific community is also demonstrated by publication records. The RDC-IAB literature database not only includes dataset descriptions and methodology reports but also lists the publications by researchers using the data offered by the RDC-IAB (see http://fdz.iab.de/en/FDZ_Publications/FDZ_Literature_Database.aspx). By 2015, the list contained 85 publications in journals, most of which (68) are in renowned refereed journals; there were also 90 monographs, 61 working papers/discussion papers, and 13 articles in collected editions (previous statistics are available in Bender et al. 2011). Note that it is likely that the number of publications is underreported because, unfortunately, not all users comply with the obligation to inform the RDC-IAB about their publications if RDC-IAB data sources are used.
5 Innovative Infrastructure for Data Access

The RDC-IAB works on its infrastructure continuously to improve the ways in which data can be accessed and processed. One of the most important infrastructure projects was the recent implementation of a remote data submission environment at the RDC. Most user projects work with weakly anonymized data, access to which is restricted to on-site use and the remote execution of data processing programs. The growing number of users increased the number of jobs submitted for remote executions and disclosure review for on-site use. In 2014, for instance, the RDC-IAB reached a total of approximately 1800 remote jobs. As a result, the RDC-IAB reached its own capacity limits more frequently. Therefore, it was necessary to improve the infrastructure and thereby manage remote data execution and the disclosure review service in a more automated way. The RDC-IAB decided to implement the Job Submission Application (JoSuA) environment in 2015 (Eberle et al. 2017). The software is maintained by the Institute of Labor Economics (IZA). The main innovation designed for JoSuA is to provide separate modes of job submission. The motivation is that typical research work requires a lot of data processing and testing for project-internal purposes only. Consequently, the bulk of output of these procedures is not suitable for use in a publication or presentation, which means that it is not necessary to export most of the output. Therefore, JoSuA provides two kinds of job submission. The first mode, “Internal Use,” is completely automated and should be used to prepare the data and test the empirical methods. In this mode, the manual output controls are replaced by a script-based automated disclosure review. The initial text output files are converted into image files and can only be previewed in JoSuA; downloading is not possible. The second mode, “Presentation/Publication,” should be used after data preparation and testing are finished. In this mode, do-files and output files are manually reviewed for disclosure risk by scientific staff members of the RDC-IAB and are subsequently made available for download via JoSuA.

Once a new job is submitted in “Publication/Presentation” mode, the previous “Internal Use” jobs are inaccessible.

The main advantages of these distinct job submission modes are obvious:

- The environment increases data security and minimizes the risk of disclosure.
- The results in “Internal Use” can be directly checked by the researcher without significant delay.
- Only output required for a presentation or publication is given to the researcher.
- All other output remains within the secure computing environment of the RDC.

Figure 4 shows the number of jobs for both modes since JoSuA was made available to RDC-IAB users. On the one hand, the number of jobs has increased enormously. In 2014, the number of jobs per month was less than 170. This number has since almost quadrupled in a typical month. On the other hand, due to JoSuA, it was possible to limit the volume of output to be checked for any disclosure risks by a scientific staff member of the RDC-IAB, since the jobs in
“Presentation/Publication” mode were only a fraction of the total. Hence, with equal personnel resources, a more intensive use of the datasets was made possible. However, due to the increase in general demand, the capacity limits of the RDC-IAB have again been reached.

A further project aimed to improve searching on the RDC-IAB datasets. To this end, a metadata management and web information system using the DDI standard were implemented. Both applications are currently loaded with the relevant contents.

Furthermore, future infrastructure development includes an enhancement and extension of the RDC-in-RDC IAB approach. The RDC-IAB plans further on-site locations both within and outside Europe. Enhancement of the concept would foster the mutual exchange of microdata access possibilities with partner institutions. Concurrently, the RDC-IAB will be involved in two feasibility studies to extend opportunities for data access via remote access. The first is planned in the framework of an extension of the Virtual Research Environment project, which was developed to support collaborative use of microdata in joint projects of spatially distributed research institutions. The Virtual Research Environment was used by the joint project “Reporting on socioeconomic development in Germany—soeb 3” (Forschungsverbund Sozioökonomische Berichterstattung 2017). The second feasibility study is planned jointly with the Center for Urban Science and Progress (CUSP) at New York University.

Finally, the RDC-IAB plans to elaborate on concepts for creating data access for linked microdatasets that cannot be stored at one of the data providers involved because of data protection, ownership claims, or other restrictions. One technical solution could involve linking the datasets held by two or more data providers and analyzing them on an encapsulated high-security server of a data custodian “on the fly” (or in the cloud) via remote access. The linkage of the datasets according to
such a concept is only temporary in the workspace of the server and remains in place only as long as needed for the statistical analyses to be completed.

6 Research Activities

A general principle of the RDC-IAB is that its staff members should not be engaged exclusively in data provision services but that they should also conduct their own research, at least to a limited extent. The idea behind this is that working on new data products and data quality requires research experience. Finally, research experience is helpful for better individual data counseling.

Many of the research activities of the RDC’s staff members are based on collaborations with national and international external researchers (e.g. Card et al. 2013; Hirsch et al. 2014; Bender et al. 2016; Fackler et al. 2016). This guarantees an intensive exchange of knowledge on new econometric and statistical methods, new trends in data handling, or the improvement of data collection and survey techniques. In addition, the RDC-IAB requires that its research activities be presented at national and international workshops and conferences. Furthermore, the RDC-IAB is involved in numerous external projects that are funded by the Germany Research Foundation, the Federal Ministry of Education and Research, or the Ministry of Labor and Social Affairs, for instance. These projects are frequently carried out in collaborations with universities, other research institutes, or research data centers. All projects are described in detail on the website of the RDC-IAB.

7 Conclusions

This paper describes the opening of German administrative labor market microdata to the international scientific community as a great success story. In 2004, the Research Data Centre at the IAB was established to offer rich administrative microdata samples of integrated labor market biographies, linked employer–employee information, and other data to a network of international scholars. All sides have profited from abandoning the closed strategy for social security data that prevailed in the past. The German Federal Employment Agency, the Ministry of Labor and Social Affairs, and other stakeholders of the IAB have benefited from improved evidence-based policy advice, the international scientific community from new opportunities to answer relevant labor market research questions with reliable and comprehensive data, and, last but not least, the IAB itself through a number of joint projects and collaboration with international researchers. An important point is that access to the data complies with strict German data protection rules. To this end, the

\[\text{See } \url{http://fdz.iab.de/en/FDZ_Projects/projects.aspx/Bereichnummer/17}.\]
Research Data Centre at the IAB has not only improved anonymization techniques but also established a data access infrastructure that meets the demanding requirements. An example is an environment that allows remote data processing. Future developments include mutual microdata exchange between partner institutions and improvements in data linkage techniques in conformity with data protection rules.

Acknowledgements For helpful comments and discussions, the authors would like to thank Daniela Hochfellner and Alexandra Schmucker.

References

Aretz B, Arntz M, Gregory T (2013) The minimum wage affects them all: evidence on employment spillovers in the roofing sector. Ger Econ Rev 14(3):282–315. https://doi.org/10.1111/geer.12012

Bellmann L (2002) Das IAB-Betriebspanel: Konzeption und Anwendungsbereiche. Allg Stat Arch 86(2):177–188

Bellmann L, Bender S, Bossler M et al (2015) LPP – Linked Personnel Panel – quality of work and economic success: longitudinal study in German establishments (data collection on the first wave). FDZ-Methodenreport 05/2015. http://doku.iab.de/fdz/reporte/2015/MR_05-15_EN.pdf. Accessed 14 Sept 2017

Bender S, Heining (2011) The Research-Data-Centre in Research-Data-Centre approach: a first step towards decentralised international data sharing. FDZ-Methodenreport 07/2011. http://doku.iab.de/fdz/reporte/2011/MR_07-11_EN.pdf. Accessed 14 Sept 2017

Bender S, Dieterich I, Hartmann B et al (2011) FDZ-Jahresbericht 2009/2010. FDZ-Methodenreport 06/2011. http://doku.iab.de/fdz/reporte/2011/MR_06-11.pdf. Accessed 14 Sept 2017

Bender S, Bloom N, Card D et al (2016) Management practices, workforce selection and productivity. CEP Discussion Paper No 1416. Centre for Economic Performance, London

Broszeit S, Laible M-C (2016) German Management Organizational Practices survey (GMOP 0813): data collection. FDZ-Methodenreport 06/2016. http://doku.iab.de/fdz/reorte/2016/MR_06-16_EN.pdf. Accessed 14 Sept 2017

Card D, Heining J, Kline P (2013) Workplace heterogeneity and the rise of West German wage inequality. Q J Econ 128(3):967–1015

Dorner M, Heining J, Jacobebbinghaus P et al (2010) The sample of integrated labour market biographies. Schmollers Jahrbuch 130(4):599–608

Dustmann C, Ludsteck J, Schönberg U (2009) Revisiting the German wage structure. Q J Econ 124(2):843–881

Eberle J, Schmucker A (2017) The establishment history panel: redesign and update 2016. Jahrb Natl Okon Stat. https://doi.org/10.1515/jbnst-2016-1001

Eberle J, Müller D, Heining J (2017) A modern job submission application to access IABs confidential administrative and survey research data. FDZ-Methodenreport 01/2017. http://doku.iab.de/fdz/reporte/2017/MR_01-17_EN.pdf. Accessed 14 Sept 2017

Ellguth P, Kohaut S, Möller I (2014) The IAB establishment panel: methodological essentials and data quality. J Labour Market Res 47(1/2):27–41

Fackler D, Schnabel C, Schmucker A (2016) Spinoffs in Germany: characteristics, survival, and the role of their parents. Small Bus Econ 46(1):93–114

Fischer G, Janik F, Müller D et al (2009) The IAB establishment panel: things users should know. Schmollers Jahrbuch 129(1):133–148
Dana Müller is head of the Research Data Center (FDZ) of the Federal Employment Agency at the Institute for Employment Research since October 2016. She studied Sociology at the Chemnitz University of Technology and worked as a researcher at the FDZ.
Joachim Möller  Studies of Philosophy and Economics at the Universities of Tübingen, Strasbourg and Konstanz. Doctorate: University of Konstanz, 1981 and Habilitation 1990. Current Position: Full Professor of Economics at the University of Regensburg, Director of the Institute for Employment Research of the Federal Employment Agency (IAB).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.