

The German Integrated Employment Biographies Sample IEBS

By Peter Jacobebbinghaus and Stefan Seth

1. Introduction

The main tasks of the German Federal Employment Agency (FEA) are job placement, the payment of unemployment benefits and the implementation of active labour market policies. To perform these tasks the FEA needs to collect data on its clients. While these data are collected for administrative purposes, they have also gained importance as a (data)base for scientific research.

With financial support of the Federal Ministry of Education and Research, the FEA has, in 2004, established a Research Data Centre (FDZ) to facilitate access to and promote labour market research based on its data (Kohlmann, 2005). The Integrated Employment Biographies Sample (IEBS) is an important step in that direction. Whereas the datasets that were made available in the past only covered employment and unemployment benefits, the IEBS also includes information on job search and active labour market programmes, making it the most comprehensive dataset regarding individual employment histories offered by the FDZ. The IEBS is described in detail in Hummel et al.¹ In the following, we give an outline of the IEBS's features.

2. Information in the Data

What's in the data: The IEBS comprises information on four fields of interest:

- Employment (12,594,862 spells).
- Unemployment benefits (2,388,627 spells).
- Active labour market policies (238,232 spells).
- Job search (1,828,266 spells).

The IEBS comes in spells. For each spell, the data contain a range of socio-economic characteristics such as age, gender, education. For employment

¹ The FDZ website, <http://fdz.iab.de>, also offers extensive information about our data; English documentation will be available at the end of 2006.

spells, the data provide detailed daily information on the employment status; additionally, information about the employing firm is available. Furthermore, the IEBS has records, on a daily basis, of received unemployment benefits, job search and participation in active labour market programmes.

The study population consists of all employees the German social insurance agencies know of, plus all individuals in contact with the FEA either because they receive unemployment benefits or they search a job. Data on employment originate from a compulsory notification scheme which requires employers to report data on employees on a yearly basis. No information is available on times of employment not liable to social security (e. g. civil servants or self-employed; marginal part-time workers, however, are included since 1999). About 80 % of the labour force are covered. Data on unemployment benefits, active labour market policies and job search are, as mentioned above, a by-product of the FEA's activities.

Sample Design: The IEBS was constructed by selecting people born on one of eight designated days of the year from the sample frame, making the IEBS a representative, 2.2-percent sample of the target population. The birthdays are kept fix, so that it is possible to track individuals' working lives – unless people get self-employed, become civil servants or take childcare leaves, etc.

Time: Data on employment and on unemployment benefits are available as far back as 1990, reliable information on active labour market programmes and job search from 2000 onwards. The FDZ plans to regularly update the IEBS. In the current version, IEBS 1.0, you can find data up to the year 2004 with the exception of employment data which stretch only until 2003. This time lag of about one year is due to the notification process and will persist in future versions of the IEBS.

Groups of variables: Here is a list of the major groups of variables to be found in the IEBS. In this survey, we focus on the most important variables. For a detailed description of all variables, please visit our website.

- IDs: person-ID, establishment-ID.
- Employment status (before, during and after the spell).
- Personal characteristics: year of birth, sex, citizenship, school education, training, severe disability status.
- Employment and job search: occupational status/working time, occupation, daily pay, industrial sector, beginning and duration of unemployment.
- Place of residence and place of work.

An essential variable is *employment status during the spell*. For employment spells, it indicates the type of employment: standard employment, apprenticeship, marginal part-time employment, etc. For unemployment benefit spells, it

shows the type of benefit received and for job search spells, the labour market status, that is, whether or not the person who searches for a job is currently employed. For policy spells, the variable signifies the type of active labour market programme. Table 1 gives an idea of how active labour market programmes are represented in the data.

Table 1
Frequency of Active Labour Market Spells

	2000	2001	2002	2003	2004
Subsidies					
Regular labour market	2,197	3,716	4,227	4,171	3,061
Second labour market	4,083	5,726	4,429	3,179	2,580
Self-Employment	1,006	1,290	1,605	2,822	4,958
Training					
Short Term	998	970	1,118	1,686	1,852
Medium & Long Term	5,131	7,158	7,503	5,507	3,505
German Courses	647	534	486	389	335
Total	14,062	19,394	19,368	17,754	16,291

Numbers refer to spells that include June, 30th.

An important feature of the IEBS is that parallel spells may occur: people might have several jobs at the same time, or receive employment benefits while on job search, or take part in certain active labour market programmes triggering both an employment spell and a ‘policy’ spell, etc. Some ‘technical’ variables were generated to facilitate data handling. Also, spells were split up, i. e. artificially multiplied in such a way that spells do not overlap.

Figure 1 illustrates a sample employment history. At first, the person is observed as employed. Towards the end of her job she registers with the FEA. She claims unemployment benefits for one year and stops searching a few weeks after the benefit has run out. We do not observe her employment status until she starts to participate in an active labour market programme – which is an employment subsidy that implies parallel employment and job search spells. The employment spell ends a few weeks after the subsidy has run out and a new job search spell begins.

Data quality: Generally, the data can be said to be very reliable (Biewen, Fitzenberger, Osikominu, and Waller, 2006). In particular, this applies to information collected not solely for statistical purposes: for instance, the data on remuneration are used by the German Statutory Pension Insurance to calculate pension claims. As another example, information on education is solid in job

seeking spells as it is relevant for job placement and the person responsible for data input has an incentive to ensure correctness; however, the educational variable is less reliable for employment spells because an incorrect information there neither hurts the employer (who gives the information) nor the employee.

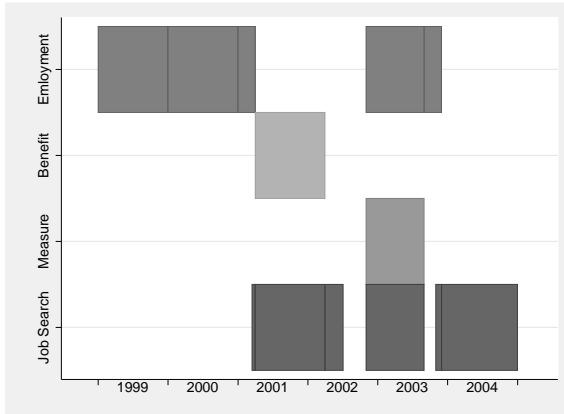


Figure 1: A Sample Employment History

The IEBS contains information from various data sources. The merging of these sources exhibits inconsistencies, most of which cannot be clarified with certainty, because the data are ambiguous. On the one hand, this underlines an advantage of the IEBS as data problems are revealed that would remain hidden otherwise. On the other hand, this means that researchers might have to throw in some pre-research effort to make the data ready for analysis. The amount of effort depends on which data sources and variables must be used to achieve the research goal; however, it can be quite substantial. Of course, the FDZ has documented known data problems and has carried out data preparation steps that doubtlessly improve the data.² Also, the FDZ offers help by providing code samples and improved variables.

3. Comparison with Related Data

Three datasets of the FDZ contain information similar to the IEBS but have a different focus: The IAB Employment Sample (IABS), the Employment Panel of the FEA (BAP) and the Linked Employer-Employee Data of the IAB

² Please note that the IEBS is being developed; future versions – including the next one, available at the end of 2006 – will be more user-friendly with respect to data quality.

(LIAB). Introductory descriptions of these datasets can be found in former editions of European Data Watch (IABS: Bender/Haas/Klose, 2000; BAP: Koch/Meinken, 2004; LIAB: Alda/Bender/Gartner, 2005). Detailed and up-to-date information is available on our website.

The current version of the IABS contains employment and unemployment benefit spells for the years 1975 to 2001. The sample is drawn as a 2 percent random sample of all individuals who were subject to social insurance contributions at least once during this period. Unemployment benefit spells are added for the individuals in the sample, which means that the IABS is representative regarding employment but – in contrast to the IEBS – not regarding benefit reception. Furthermore, the IABS contains no information on job search and participation in labour market programmes. On the other hand, it covers a longer time span and, due to more rigorous data preparation, shows fewer inconsistencies.

In contrast to IEBS and IABS, BAP is a panel study. The sample is drawn, at the end of each quarter, by selecting those born on one of 7 designated days from all individuals who are employed on the reference date – thus creating panel data without attrition. In that the sample is based on birthdays, BAP equals IEBS, in that the target population is employees, BAP resembles IABS. In addition to individual employment data (wage, profession, gender, etc.), BAP contains information about the employing firm. At the moment, the data range from 1998 to 2003; after the next update, scheduled for summer 2006, BAP's data will stretch up to 2005. Besides its up-to-dateness, the main advantage of BAP compared with IEBS or IABS is that it is easier to analyse because it contains panel data rather than spell data.

The LIAB are linked employer-employee data. Individual information regarding employment and unemployment benefit are linked to the firm survey IAB Establishment Panel. For the firms in the sample information of all employees subject to social insurance contributions in the firm are added. Cross sectional and longitudinal versions are available. The cross sectional versions include more firms and more years but only employee information on the day of the firm questionnaire. The longitudinal versions include less firms but the employment histories of their employees for a longer period.

4. Research Topics

The IEBS includes detailed information on individual employment histories and provides a rich set of covariates. This allows a wide range of research questions to be addressed. These topics include wages (rigidities, discrimination, dispersion, effects of education or union membership, etc.), employment and unemployment duration, and mobility (regional, occupational). Numer-

ous studies on these topics are based on the IABS.³ With the availability of additional information in the IEBS these studies can be enhanced. Data on job search allow to reconsider the definition of unemployment in duration analysis. Data on active labour market policies allow the distinction between regular and subsidised employment in wage analysis.

More importantly, the inclusion of information on active labour market policies extends the range of research topics, especially to evaluation studies. Biewen et al. (2006) provide an interesting example. They estimate, with IEB data,⁴ the employment effects of short and medium term further training and show that the effectiveness of these measures varies considerably among sub-populations, i. e. it depends on gender, region and duration of unemployment. IEB data are the basis for many other evaluation studies related to the so-called Hartz-reforms.⁵

5. Data Access

The FDZ offers access to the IEBS in two ways: via remote data access and via research stays at the FDZ in Nuremberg. Remote data access means that the researcher uses the data documentation and test data to prepare statistical code and sends it to the FDZ by e-mail.⁶ The FDZ staff runs the code, checks the files that contain the results and erases information suited to identify individuals. The remaining results are send to the researcher. Alternatively, direct data access is possible via research stays at the FDZ in Nuremberg. Regarding the results, the rules described above apply, i. e. information suited to identify individuals is erased.

Before data access is possible in either way, an application has to be filed by the researcher, approved by the Federal Ministry of Labour and Social Affairs, and a contract has to be signed. The legal background of providing data access to the IEBS is § 75 book 10 of the German Social Code (*SGB*). All person-specific information collected by the FEA in order to provide unemployment insurance, measures of labour market policy and job placement are so-called social data (*Sozialdaten*) for which specific confidentiality rules apply. Scientific use of social data requires the following conditions to be met:

³ An up-to-date list of all publications based on the FDZ datasets that are reported to us on our website.

⁴ The data used in this study are not identical to the IEBS offered at the FDZ, but very similar.

⁵ The Hartz-reforms changed old and introduced new labour market instruments in Germany. For a selection of these studies, see Brinkmann/Hujer/Koch (2006).

⁶ We currently accept code written in STAPA, SPSS, and SAS but plan to include further statistical software packages. Please visit website to check the current terms.

- Scientific research regarding social security.
- Prevailing public interest.
- Permission of the Federal Ministry of Labour and Social Affairs.

The FDZ coordinates the application process which normally takes less than two weeks. Forms and guidance can be found on our website.

6. Outlook

The IEBS will be updated on a regular basis. The inclusion of more recent data, especially of more recent employment spells, will further improve the possibility to study employment effects of policy measures. In addition to the extension of the time covered, further information will be included on new policy measures like employment in personal service agencies. Unemployment benefit spells will also include information on the unemployment benefit based on Social Code II (*ALG II*) which was introduced in 2005. Along with this new, means tested benefit comes valuable information about the household context, more precisely, on all household members obligated to financially support each other (*Bedarfgemeinschaft*).

The main goal of providing the IEBS is to make well-documented policy data accessible to the scientific community by means of a standardised dataset. While this goal has already been achieved to a great extent, the FDZ would like to broaden the ways in which IEBS data may be accessed. Currently, the IAB is working on a concept of factual anonymisation of IEB data. If factual anonymisation proves feasible, this concept will be implemented and a scientific use file of the IEBS offered by the FDZ.

References

- Alda, H./Bender, S./Gartner, H.* (2005): The linked employer-employee dataset created from the IAB establishment panel and the process-produced data of the IAB (LIAB), *Schmollers Jahrbuch* 125, 327–336.
- Bender, S./Haas, A./Klose, C.* (2000): The IAB Employment Subsample 1975–1995, *Schmollers Jahrbuch* 120, 649–662.
- Biewen, M./Fitzenberger, B./Osikominu, A./Waller, M.* (2006): Employment Effects of Short and Medium Term Further Training Programs in Germany in the Early 2000s, unpublished paper.
- Brinkmann, C./Hujer, R./Koch, S.* (2006): Evaluation aktiver Arbeitsmarktpolitik in Deutschland, forthcoming in *Journal for Labour Market Research (ZAF)*.

- Hummel, E./Jacobebbinghaus, P./Kohlmann, A./Oertel, M./Wübbecke, C./Ziegerer, M.* (2005): Stichprobe der Integrierten Erwerbsbiografien – IEBS 1.0, FDZ Datenreport, 6.
- Koch, I./Meinken, H.* (2004): The Employment Panel of the German Federal Employment Agency, *Schmollers Jahrbuch* 124, 315–325.
- Kohlmann, A.* (2005): The Research Data Centre of the Federal Employment Service in the Institute for Employment Research, *Schmollers Jahrbuch* 125 (3), 437–447.