Linking LEEP-B3 Survey Data with Administrative IAB Data

Peter Jacobebbinghaus
Stefan Seth
Martin Diewald
Reinhard Schunck
Anja-Kristin Abendroth
Silvia Maja Melzer
Stephanie Pausch
Mareike Reimann
Björn Andernach
Linking LEEP-B3 Survey Data with Administrative IAB Data

SFB 882 Technical Report, No. 21
DFG Research Center (SFB) 882 From Heterogeneities to Inequalities
Project “Interactions Between Capabilities in Work and Private Life: A Study of Employees in Different Work Organizations” (B3)
Bielefeld, May 2015

SFB 882 Technical Report Series
General Editors: Stefan Liebig and Johanna Vompras
ISSN 2196-985X

This publication has been funded by the German Research Foundation (DFG).

SFB 882 Technical Reports are refereed scholarly papers. Submissions are reviewed by peers in a two-stage SFB 882 internal and external refereeing process before a final decision on publication is made.

The Technical Report Series is a forum for presenting technical works (e.g. data documentation, field reports etc.) in progress. Readers should communicate comments on the manuscript directly to the author(s).

The papers can be downloaded from the SFB 882 website http://www.sfb882.uni-bielefeld.de/

SFB 882 “From Heterogeneities to Inequalities”
University of Bielefeld
Faculty of Sociology
PO Box 100131
D-33501 Bielefeld
Germany
Phone: +49-(0)521-106-4942 or +49-(0)521-106-4613
Email: office.sfb882@uni-bielefeld.de
Web: http://www.sfb882.uni-bielefeld.de/
Content

1 Introduction ......................................................................................................................................... 3
2 Schematic Overview .......................................................................................................................... 4
3 Description of Data Components ................................................................................................... 7
   3.1 Establishment Survey .................................................................................................................. 7
   3.2 Employee and Partner Surveys .................................................................................................. 7
   3.3 Extracts from the Integrated Employment Biographies (IEB) ................................................. 7
      3.3.1 IEB for Work Organizations ............................................................................................ 8
      3.3.2 IEB for Persons .................................................................................................................. 10
   3.4 Extracts from the Employment History Panel (BHP) ............................................................ 11
      3.4.1 BHP for Work Organizations ............................................................................................ 12
      3.4.2 BHP for Persons .................................................................................................................. 13
4 Record Linkage of Partner Survey Data ......................................................................................... 15
5 Analysis of Potentials ..................................................................................................................... 17
6 Future Developments .................................................................................................................... 18
References ........................................................................................................................................... 19
Appendix 1: List of IEB variables ...................................................................................................... 20
Appendix 2: List of BHP characteristics ............................................................................................. 23
1 Introduction

The project B3 “Interactions Between Capabilities in Work and Private Life: A Study of Employees in Different Work Organizations” is part of the Collaborative Research Center’s program “From Heterogeneities to Inequalities” (SFB 882), which is supported by a grant from the German Research Foundation (DFG). Project B3 examines the role of the workplace context in producing social inequalities. It considers the occupational and private situation of employees and focuses on how these life areas influence each other.

In order to create an extensive database, an employer survey of 115 work organizations was carried out in 2012 (see Pausch et al. 2014). A random sample of employees was then drawn from 100 of the surveyed companies. These employees and, if possible, their partners were interviewed by telephone (see Abendroth et al. 2014; Reimann et al. 2014). With the consent of those interviewed, data from the employee and partner surveys were linked with the employer survey. The result was a linked employer–employee data set that enabled us to include the circumstances of the workplace and the employees’ private circumstances in the analysis of the employees’ situation.

If the respondents agreed to have their data linked, their information was supplemented by administrative data from the German Institute for Employment Research (IAB). These administrative data provide information about individual employment histories and about companies they were employed at. Thus, the data set of the B3 Linked Employer–Employee Panel (LEEP-B3) consists of three components: the employer survey, the employee and partner surveys and the IAB administrative data.

This technical report provides an overview of the complex data structure of the LEEP-B3 data. In addition, extracts from the administrative IAB data are described. Detailed documentation of the employer survey and of the employee and partner surveys can be found in Pausch et al. (2014) and in Abendroth et al. (2014).

This report is structured as follows: Section 2 presents a schematic overview of the individual data components and the possibilities for linking the survey results. Section 3 describes the data components in more detail, focusing in particular on the selection criteria and the contents of the administrative data. Section 4 outlines the procedure for the Record Linkage of the partner survey. Section 5 provides some examples to illustrate the analytical potentials resulting from the interaction of the individual data components, and Section 6, the final section, provides a discussion of the subsequent survey waves.
2 Schematic Overview

The specific components of the LEEP-B3 data set are shown in Figure 1. The survey data form the core, consisting of the establishment survey and of the employee and partner surveys, which were conducted after the establishment survey. For the establishment survey, interviews with 115 work organizations were conducted in 2012 (see Pausch et al. 2014). After the employer survey was completed, the employee and partner surveys were carried out between the fall of 2012 and the spring of 2013 (see Abendroth et al. 2014). These surveys were then linked to create a data set in a linked employer–employee design. However, the information obtained from the employee interviews could only be assigned to the administrative data if the respondents had agreed for their data to be linked. Of the 6,454 employees surveyed, 5,653 (88%) agreed to the linkage; for the remaining 12 percent, linkage was not possible.

Of the main respondents, 5,368 indicated that they had a partner, and 2,185 of these partners were also interviewed. This additional information allowed a more detailed investigation of the employees’ private lives. The partners were also asked whether their information could be linked with the other IAB data; 1,888 of the partners consented to such a linkage and provided their names, dates of birth, and addresses, on which the linking process is based.

The survey data can be supplemented with administrative data from IAB that originate from two sources: the Integrated Employment Biographies (IEB), which contain individual data about employment history, and the Establishment History Panel (BHP), which includes information about the work organizations’ employment structures. The IEB and the BHP data are collected since 1975. From these administrative data, two types of extracts were processed:

- Extracts for supplementing information about the interviewed work organizations (IEB for Work Organizations and BHP for Work Organizations, shown on the left side of Fig. 1) and

- Extracts for supplementing information about the interviewed individuals (IEB for Persons and BHP for Persons, shown on the right side of Fig. 1).

Both types of extracts contain largely the same variables, but they differ in the selection of the individuals and work organizations included.
IEB stands for Integrated Employment Histories. The IEB contains episode data on employment, benefit receipt, participation in measures and job search.

The available data status of the IEB for the creation of the first data version was from 2011. Throughout the project duration more recent data are supposed to be added.

BHP stands for Establishment History Panel. The BHP primarily contains information about the employee and wage structure of the establishments.

The gray ellipses represent via which establishment or personal numbers, the data extracts can be linked. Different establishment and personal numbers were assigned in order to prevent data links, in the cases where respondents did not agree. Accordingly, in the employee and partner survey the betidL is only included for the 5,653 respondents, who agreed to be linked with other data.
The 

The extract *BHP for work organizations* contains information about the 115 interviewed establishments that was collected annually for the period 1975–2011. For the extract *IEB for Work Organizations*, we selected the employment biographies of all individuals who were employed for at least one day at one of the 115 interviewed work organization since January 1, 2007. On the basis of these supplementary administrative data for work organizations, we were able to trace the development of the establishments. In particular, the individual data from the extract *IEB for Work Organizations* allowed us to trace detailed changes in labor forces and wage structures since 2007.

The extracts of supplemental data on interviewed individuals (*IEB for Persons* and *BHP for Persons*) contain information about individuals who agreed for their data to be linked with other data from IAB. Since the sampling of the employee survey was based on IAB data, these employees were selected on the basis of personal numbers in the IEB. For the partners, these personal numbers had to be determined by a record-linking method, which was based on the names, addresses, and dates of birth provided by the respondents for that purpose. This allowed us to identify 1,378 of the 1,888 partners who agreed for their data to be linked with the administrative data from IAB (see Sect. 4). There were 21 individuals who participated in both the employee survey and the partner survey. Thus, the extract *IEB for Persons* includes the employment histories of 7,010 individuals which can be linked with the employee and partner surveys. The extract *BHP for Persons* complements these employment histories with information about the establishments where such individuals were previously employed.

Linkage of the six described components of the LEEP-B3 data was carried out by means of personal and establishment identification numbers. The gray ovals in Figure 1 indicate which components can be linked by use of which variables. Different person and establishment numbers were assigned to prevent linkage of data for which respondents’ consent had not been obtained.

The following sections describe the contents of the LEEP-B3 data components in more detail.
3 Description of Data Components

The LEEP-B3 data include data from the establishment survey and from the employee and partner surveys, and administrative data that supplement the survey data with other characteristics and periods of time. Only a brief description of the survey data will be included here because these data have been documented in separate technical reports. The administrative data, however, will be described in more detail.

3.1 Establishment Survey

The population of the establishment survey consisted of workplaces across almost all economic sectors that had more than 500 employees who were subject to social insurance contributions on December 31, 2010. The drawing of the gross sample was based on the IAB Employment History, which contains all German companies with at least one employee paying social insurance contributions. The gross sample of 539 randomly sampled work organizations yielded 115 employer interviews. Issues covered by this survey included employment structure, employment policy measures, equal opportunities, work–life balance, and health. A comprehensive and detailed description of the establishment survey is provided in the technical report by Pausch et al. (2014).

3.2 Employee and Partner Surveys

From 100 of the 115 interviewed work organizations, a random sample was drawn to arrange for telephone interviews with employees who were employed on December 31, 2011. The employees’ addresses were submitted to the survey institute,¹ which interviewed 6,454 employees (Abendroth et al., 2014), of whom 5,368 indicated that they had a partner at the time of the survey. These employees were asked to provide their partner’s contact information. Interviews, which were slightly shorter than the employee interviews, were then conducted with 2,185 partners who had consented to participate. Issues covered by the employee and partner surveys included occupation, private life, work–life balance, health, satisfaction, and social demographics.

3.3 Extracts from the Integrated Employment Biographies (IEB)

The Integrated Employment Biographies (IEB) contains information on employment (since 1975), benefits received (since 1975), unemployment and job searches (since 2000), and employment policy measures (since 2000). Information from different data sources of the Federal Employment Agency are merged in the IEB, preserving the daily accuracy of the data. Data for the new federal states were first collected in 1990 and can be regarded as complete since 1993. A detailed description of the administrative personal data is provided in

¹ The Institute for Social Research and Communication (SOKO), Bielefeld.

The LEEP-B3 data include two IEB extracts: the extract for establishments and the extract for employees (see Fig. 1). For the vast number of characteristics, the variable names and their values in the IEB extracts match those in the Sample of Integrated Labor Market Biographies 1975–2010 (SIAB 7510). These characteristics are listed in Appendix 1. A more detailed description can be found in Berge, König and Seth (2013). Additional characteristics were specifically generated for the LEEP-B3 data, as described in Appendix 1.

The following sections describe the structures of the two IEB extracts:

- The IEB Extract for Work Organizations contains the employment histories of all individuals who have been employed for at least one day by one of the 115 surveyed work organizations since January 1, 2007. It serves to describe the interviewed establishments.

- The IEB Extract for Persons includes the employment histories of all interviewed individuals who agreed for their data to be linked with other IAB data. It is used to supplement the personal survey with employment-biographical information from administrative sources.

3.3.1 IEB for Work Organizations

The extract IEB for Work Organizations is a set of personal data that serves to characterize the workforce of the 115 interviewed establishments. It contains the employment histories of all individuals who have been employed for at least one day in one of the 115 work organizations since January 1, 2007. Since information on all employees of the establishments is available in the extract, workplace-specific figures can be calculated. For example, it is possible to generate indicators for occupational segregation by gender or nationality, or any indices of occupational wage inequality. By linking these ratios with the employer and employee surveys, it is possible to analyze a variety of characteristics.

The extract IEB for Work Organizations contains information on 274,576 individuals, 232,955 of whom were employed in one of the 115 work organizations that took part in the employee surveys. Not all the personal employment histories date back to 1975. Table 1 shows the number of individuals for whom information was available on June 30 for the years selected. Of the 28,994 individuals included in 1975, 4,175 were already employed in one of the 115 responding establishments in 1975.
Table 1. Number of individuals and establishments in the IEB Extract for Work Organizations for selected years

<table>
<thead>
<tr>
<th>Year</th>
<th>Individuals</th>
<th>Employed in one of the 115 work organizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>28,994</td>
<td>4,175</td>
</tr>
<tr>
<td>1980</td>
<td>48,182</td>
<td>9,568</td>
</tr>
<tr>
<td>1985</td>
<td>67,929</td>
<td>14,651</td>
</tr>
<tr>
<td>1990</td>
<td>92,237</td>
<td>24,143</td>
</tr>
<tr>
<td>1995</td>
<td>128,910</td>
<td>41,345</td>
</tr>
<tr>
<td>2000</td>
<td>172,305</td>
<td>78,251</td>
</tr>
<tr>
<td>2005</td>
<td>206,591</td>
<td>121,866</td>
</tr>
<tr>
<td>2010</td>
<td>230,945</td>
<td>157,869</td>
</tr>
<tr>
<td>2011</td>
<td>227,524</td>
<td>157,676</td>
</tr>
</tbody>
</table>

Table 2 shows the number of data rows according to data sources on June 30 for the selected years. Because episodes often overlap, it appears that the number of data rows is significantly larger than the number of individuals; one reason is that many individuals who are receiving unemployment benefits are also looking for new employment. The table shows that information on employment histories and unemployment benefit histories dates back to 1975, while the other sources were provided only from 2000 or 2005 onward.

Table 2. Number of data rows by data source

<table>
<thead>
<tr>
<th>Year</th>
<th>BeH</th>
<th>LeH</th>
<th>LHG</th>
<th>ASU</th>
<th>MTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>29,134</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>1980</td>
<td>47,382</td>
<td>1,135</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>1985</td>
<td>65,318</td>
<td>2,909</td>
<td>0</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>1990</td>
<td>92,296</td>
<td>3,629</td>
<td>0</td>
<td>35</td>
<td>9</td>
</tr>
<tr>
<td>1995</td>
<td>125,908</td>
<td>5,537</td>
<td>0</td>
<td>900</td>
<td>118</td>
</tr>
<tr>
<td>2000</td>
<td>172,101</td>
<td>5,864</td>
<td>0</td>
<td>9,033</td>
<td>3,133</td>
</tr>
<tr>
<td>2005</td>
<td>203,570</td>
<td>4,347</td>
<td>11,265</td>
<td>20,674</td>
<td>3,322</td>
</tr>
<tr>
<td>2010</td>
<td>234,295</td>
<td>5,117</td>
<td>11,484</td>
<td>23,948</td>
<td>3,513</td>
</tr>
<tr>
<td>2011</td>
<td>231,976</td>
<td>4,765</td>
<td>10,088</td>
<td>20,904</td>
<td>2,449</td>
</tr>
</tbody>
</table>

Notes: BeH = employment history, LeH = benefit receipt history, LHG = unemployment benefit II recipient histories, ASU = jobseeker history, MTH = participants-in-measures history file.
3.3.3 IEB for Persons

The extract *IEB for Persons* is a set of personal data that serves to supplement the employee and partner surveys with employment-biographical information from administrative sources. This is possible for all employees and partners who have agreed for their information to be linked with other IAB data. Information about the partners was linked by means of record linkage. This was done for 5,653 main respondents and 1,378 partners; 21 individuals were both a main respondent and a partner. Thus, the extract *IEB for Persons* includes the employment histories of 7,010 individuals. A source such as this that links the survey data with the administrative data enables researchers to examine, for example, how different employment histories influence respondents’ current level satisfaction or to calculate characteristics such as period of employment and work experience, which are often needed for wage analysis.

Table 3 shows the numbers of individuals for whom information was available on June 30 of selected years. In 1975, only 8 individuals were included because the employee survey only covered individuals who were born in or after 1960. In 1975, none of the individuals was employed in one of the 100 establishments that took part in the employee survey. Table 4 shows the number of data rows for June 30 of selected years by data sources.

### Table 3. Number of employees and partners by year and source

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Main employees</th>
<th>Employed in one of the 100 establishments</th>
<th>Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>35</td>
<td>8</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>1980</td>
<td>620</td>
<td>508</td>
<td>68</td>
<td>112</td>
</tr>
<tr>
<td>1985</td>
<td>1,467</td>
<td>1,204</td>
<td>223</td>
<td>269</td>
</tr>
<tr>
<td>1990</td>
<td>2,389</td>
<td>1,968</td>
<td>520</td>
<td>428</td>
</tr>
<tr>
<td>1995</td>
<td>3,481</td>
<td>2,851</td>
<td>913</td>
<td>644</td>
</tr>
<tr>
<td>2000</td>
<td>4,818</td>
<td>3,950</td>
<td>1,560</td>
<td>880</td>
</tr>
<tr>
<td>2005</td>
<td>5,701</td>
<td>4,727</td>
<td>2,579</td>
<td>989</td>
</tr>
<tr>
<td>2010</td>
<td>6,463</td>
<td>5,420</td>
<td>4,479</td>
<td>1,062</td>
</tr>
<tr>
<td>2011</td>
<td>6,576</td>
<td>5,520</td>
<td>5,200</td>
<td>1,075</td>
</tr>
</tbody>
</table>
### Table 3. Number of data rows by data sources

<table>
<thead>
<tr>
<th>Year</th>
<th>BeH</th>
<th>LeH</th>
<th>LHG</th>
<th>ASU</th>
<th>MTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>35</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1980</td>
<td>608</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1985</td>
<td>1,410</td>
<td>61</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1990</td>
<td>2,368</td>
<td>109</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1995</td>
<td>3,389</td>
<td>155</td>
<td>0</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>2000</td>
<td>4,766</td>
<td>173</td>
<td>0</td>
<td>270</td>
<td>116</td>
</tr>
<tr>
<td>2005</td>
<td>5,577</td>
<td>148</td>
<td>218</td>
<td>592</td>
<td>126</td>
</tr>
<tr>
<td>2010</td>
<td>6,673</td>
<td>92</td>
<td>137</td>
<td>505</td>
<td>72</td>
</tr>
<tr>
<td>2011</td>
<td>6,879</td>
<td>55</td>
<td>98</td>
<td>346</td>
<td>36</td>
</tr>
</tbody>
</table>

Notes: BeH = employment history, LeH = benefit receipt history, LHG = unemployment benefit II recipient histories, ASU = jobseeker history, MTH = participants-in-measures history file.

### 3.4 Extracts from the Employment History Panel (BHP)

The Establishment History Panel (BHP) is based on employment data obtained from the social insurance reporting (social-security notifications). The employment data are aggregated based on the establishment number issued by the Federal Employment Agency at the organizational level. The BHP includes three types of work organization data. The core data set of the BHP consists of inventory data. In addition to information about the location and the economic sector of the work organizations, it provides detailed information about the composition of their workforce as of the reference date June 30 of the year selected. These employee groups are subdivided into gender, education, nationality, and age groups. Ratios for the occupational wage structure are also included. In addition to the core data set of inventory data, there are employment flows and workplace dynamics modules; the employment flows module contains information about entries and exits of employees according to different demographic characteristics, while the workplace dynamics module provides information about the founding and closure of establishments.

The BHP goes back to 1975 for West Germany and to 1991 for East Germany. The database is updated at regular intervals to include data for recent time periods. In the first extracts for LEEP-B3, the core data set and the workplace dynamics module ranged to 2011 and the employment flows module to 2010. From 1975 onward, all establishments with at least one employee who is subject to social insurance contributions for West Germany are included in the BHP. Since 1999, the BHP has also contained establishments that have only employees who are marginally employed. A list of the variables included in the LEEP-B3 extracts can be found in Appendix 2; a detailed description of the BHP is provided in Gruhl, Schmucker, and Seth (2012).

In addition to the IEB extracts, the LEEP-B3 data contain two BHP extracts, which differ in the selection of work organizations but not in the canon of variables:
The extract *BHP for Work Organizations* contains the 115 interviewed establishments. It is used to describe the development of these establishments.

The extract *BHP for Persons* includes the establishments that employed the 7,010 individuals who participated in the employee and partner surveys, agreed for their data to be linked with other IAB data, and were linked successfully. This extract supplements the personal survey with information about the types of work organizations in which the respondents have been employed throughout their work life.

### 3.4.1 BHP for Work Organizations

The extract *BHP for Work Organizations* provides administrative information to supplement the information about the 115 work organizations in the establishment survey. Accordingly, all data rows of the *BHP* that related to these 115 establishments were selected for this extract. As with the extract *IEB for Persons*, the purpose of this extract was to supplement the survey data with key parameters about companies. The selection of these key parameters is fixed in the extract *BHP for Work Organizations*; however, the values for some of the organizations date back to 1975. For example, the characteristics allow the development of a typology of growing and shrinking companies or of typologies of companies with a high or low proportion of women or foreigners.

The core data set contains 3,151 observations. Table 5 shows the number of cases for each of the selected years and the totals for West and East Germany, respectively. Berlin is included with East Germany because the BHP does not distinguish between East Berlin and West Berlin. For 64 of the 115 companies, observations date back as far as 1975. The four companies included under “East Germany” for the years 1975 through 1990 were based in Berlin.

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>West Germany</th>
<th>East Germany (including Berlin)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>64</td>
<td>60</td>
<td>4</td>
</tr>
<tr>
<td>1980</td>
<td>68</td>
<td>64</td>
<td>4</td>
</tr>
<tr>
<td>1985</td>
<td>67</td>
<td>63</td>
<td>4</td>
</tr>
<tr>
<td>1990</td>
<td>69</td>
<td>65</td>
<td>4</td>
</tr>
<tr>
<td>1995</td>
<td>86</td>
<td>72</td>
<td>14</td>
</tr>
<tr>
<td>2000</td>
<td>99</td>
<td>82</td>
<td>17</td>
</tr>
<tr>
<td>2005</td>
<td>110</td>
<td>89</td>
<td>21</td>
</tr>
<tr>
<td>2010</td>
<td>114</td>
<td>91</td>
<td>23</td>
</tr>
<tr>
<td>2011</td>
<td>115</td>
<td>92</td>
<td>23</td>
</tr>
</tbody>
</table>
Table 6 shows when the establishments were included in the BHP data for the first time. Most establishments have been included since 1975; 51 establishments were added after 1975.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>62</td>
</tr>
<tr>
<td>1976</td>
<td>2</td>
</tr>
<tr>
<td>1977</td>
<td>1</td>
</tr>
<tr>
<td>1979</td>
<td>1</td>
</tr>
<tr>
<td>1990</td>
<td>1</td>
</tr>
<tr>
<td>1991</td>
<td>1</td>
</tr>
<tr>
<td>1992</td>
<td>1</td>
</tr>
<tr>
<td>1994</td>
<td>2</td>
</tr>
<tr>
<td>1995</td>
<td>3</td>
</tr>
<tr>
<td>1996</td>
<td>1</td>
</tr>
<tr>
<td>1997</td>
<td>3</td>
</tr>
<tr>
<td>1998</td>
<td>5</td>
</tr>
<tr>
<td>1999</td>
<td>1</td>
</tr>
<tr>
<td>2000</td>
<td>2</td>
</tr>
<tr>
<td>2001</td>
<td>1</td>
</tr>
<tr>
<td>2002</td>
<td>2</td>
</tr>
<tr>
<td>2004</td>
<td>4</td>
</tr>
<tr>
<td>2005</td>
<td>4</td>
</tr>
<tr>
<td>2006</td>
<td>1</td>
</tr>
<tr>
<td>2008</td>
<td>1</td>
</tr>
<tr>
<td>2009</td>
<td>2</td>
</tr>
</tbody>
</table>

39 of the 51 establishments are included in the company dynamics module. For these establishments the variable “type of foundation” contains more information about the first occurrence of the organization in the data set.

<table>
<thead>
<tr>
<th>Type of foundation</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID change</td>
<td>4</td>
</tr>
<tr>
<td>Spin-off/pulled</td>
<td>7</td>
</tr>
<tr>
<td>Spin-off/pushed</td>
<td>2</td>
</tr>
<tr>
<td>New foundation (small)</td>
<td>11</td>
</tr>
<tr>
<td>New foundation (medium/large)</td>
<td>3</td>
</tr>
<tr>
<td>New foundation (chunky)</td>
<td>10</td>
</tr>
<tr>
<td>Unknown</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>39</td>
</tr>
</tbody>
</table>

### 3.4.2 BHP for Persons

The extract *BHP for Persons* includes the establishments in which those who responded to the employee and partner surveys were employed if the respondents had agreed for their employment. Of the 51 companies, 12 are not included in the company dynamics module. In 9 cases, this occurred in 1991 or 1992 for the first time in East Germany, so classification is not possible.
data to be linked with the IAB data and were in fact linked successfully. This extract is used to supplement the personal survey with information about the types of work organizations in which the respondents have been employed throughout their work life. Along with the employee survey, it enabled us to examine whether someone who had always been working in the same sector had expectations of his employer that were different from those of someone who had experienced different organizational contexts.

The core data set contains information about 22,269 work organizations in which 7,010 individuals who participated in the employee and partner surveys were employed in 1975. Table 8 shows the number of cases for selected years and the totals for East and West Germany, respectively, with Berlin considered a part of East Germany.

Table 8. Number and location of establishments by year

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>West Germany</th>
<th>East Germany</th>
<th>Unknown location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>7,446</td>
<td>7,220</td>
<td>226</td>
<td>0</td>
</tr>
<tr>
<td>1980</td>
<td>8,675</td>
<td>8,400</td>
<td>274</td>
<td>1</td>
</tr>
<tr>
<td>1985</td>
<td>9,781</td>
<td>9,458</td>
<td>323</td>
<td>0</td>
</tr>
<tr>
<td>1990</td>
<td>10,960</td>
<td>10,558</td>
<td>402</td>
<td>0</td>
</tr>
<tr>
<td>1995</td>
<td>13,946</td>
<td>11,604</td>
<td>2,341</td>
<td>1</td>
</tr>
<tr>
<td>2000</td>
<td>15,573</td>
<td>12,922</td>
<td>2,651</td>
<td>0</td>
</tr>
<tr>
<td>2005</td>
<td>15,158</td>
<td>12,574</td>
<td>2,584</td>
<td>0</td>
</tr>
<tr>
<td>2010</td>
<td>13,885</td>
<td>11,521</td>
<td>2,363</td>
<td>1</td>
</tr>
<tr>
<td>2011</td>
<td>13,560</td>
<td>11,268</td>
<td>2,292</td>
<td>0</td>
</tr>
</tbody>
</table>
4 Record Linkage of Partner Survey Data

Within the scope of the employee and partner surveys, 2,185 partners were interviewed, 1,888 of whom agreed for their information to be linked with IAB data. Unlike the main respondents, the partners were not assigned personal identification numbers that could be used for this purpose. For this reason, the linkage was carried out by the German Record Linkage Center on the basis of the names, addresses, and dates of birth obtained. As a result of the record linkage, the data of 1,378 of the 1,888 respondents’ partners were linked with administrative IAB data through address details. The following is a brief description of the linking procedure and of the results; a more detailed description can be found in Schild and Antoni (2014).

Partners who agreed for their information to be linked with IAB data were asked to provide their surname, given name, birth name, and date of birth. If the respondents did not live in the same household as the main respondent, they were also asked for their address (street, house number, postal code, city). In cases where the partners did not live together, the address of the main respondent was also included because the partners had most likely shared a household in the past. The address file that accompanies the IAB process data contains the names, addresses, and dates of birth of all employees subject to social insurance contributions and of all clients of the Federal Employment Agency. Individuals who did not belong to at least one of those groups were not included. This included unemployed individuals who were not registered with the Federal Employment Agency, self-employed individuals, and civil servants.

To identify the interviewed partners among the IAB data, the two address lists to be matched were uniformly processed in a first step (preprocessing). Then the exact matches were determined: first with and then without the street number. Cases for which no exact match was found were merged into blocks. Within these blocks, a probabilistic matching was performed. For this purpose, similarity measures were calculated within the blocks, and based on these measures, it was decided whether it was the same individual or not. First, block forming was carried out by using postal codes and then, for the remaining cases, by birth years. Table 9 shows the number of matches by type of match.

---

3 The linkage of the data was carried out by the German Record Linkage Center (RLC German, www.record-linkage.de), which is funded by the German Research Foundation.
4 For further details on record linkage, see Schnell (2010) and Schnell (2013).
### Tabelle 4. Results of the record linkage

<table>
<thead>
<tr>
<th>Description</th>
<th>Absolute</th>
<th>in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact match of given name, surname, date of birth, and address</td>
<td>650</td>
<td>34.4</td>
</tr>
<tr>
<td>Exact match of given name, surname, date of birth, and address without street number (often the street number was missing in the data from the Federal Employment Agency)</td>
<td>316</td>
<td>16.7</td>
</tr>
<tr>
<td>Probabilistic matching for block building by postal codes (5 digits),</td>
<td>403</td>
<td>21.3</td>
</tr>
<tr>
<td>- including 17 ambiguous cases of type A and 4 ambiguous cases of type B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probabilistic matching for block building by birth year</td>
<td>9</td>
<td>0.5</td>
</tr>
<tr>
<td>- including 4 ambiguous cases of type A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Match</td>
<td>505</td>
<td>27.0</td>
</tr>
<tr>
<td>- including 17 cases that were excluded during preprocessing because of missing names or birth dates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1,888</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Of the probabilistic matches, 25 were identified as “ambiguous cases” and labeled as such:

- **Type A (21 cases):** Match with surname, given name, and date of birth within the same location, or
- **Type B (4 cases):** Match with surname, given name, and date of birth, but with considerably different surname (fictional example: Seemaier/Zehetmayr) probably caused by oral transmission.

The type of linkage including the marking of the ambiguous cases was contained in the data (variable link-pa).

During data linkage, it became apparent that 21 individuals participated in both the employee survey and the partner survey. One reason is that the address data from IAB contained no information about existing partnerships; another is that it was not unusual for both partners to be employed in the same company, especially if it was a large one. In these cases, it was possible that both partners were chosen for the employee survey. If they then provided the contact details of their partners as part of the employee survey, they were asked for an interview a second time.
5 Analysis of Potentials

The central evaluation level of LEEP-B3 data is the employee survey. It contains extensive information on the following topics:

- Work, family
- Work–life balance
- Health
- Psychological contract
- Attitudes toward gender roles and foreigners
- Partnership quality
- Satisfaction
- Socio-demographics.

A special feature of the LEEP-B3 data is that it allows researchers to add more information to the contents of the employee survey. The following list provides an overview of which types of information are available in which data components:

- Partner information from the partner survey:
  - Employment and wages of the partner
  - Partnership quality
  - Satisfaction of the partner.

- Information about the current employer from the establishment survey
  - Competitive pressure, employment structure, payment system
  - Organizational measures to promote families, women, employees with a migration background, or health.

- Organizational information about the current employer of the extracts IEB and BHP for Establishments:
  - Employment structure of the establishment by gender, qualification, age, nationality, job tenure, parenthood, and parental leave
  - Wage level and wage structure of the establishment, high-wage vs. low-wage establishment
  - Wage differences among groups of individuals in an establishment
  - Development of the establishment over the last years.

- Information on the employment history of the main respondents and their partners from the extracts IEB and BHP for Persons:
  - Previous periods of employment and unemployment
  - Interruptions of employment due to motherhood
  - Change of occupation
  - Wage developments of respondents over time
  - Wage developments of occupational groups over time
  - Wage and employment structure of the establishments in which respondents were previously employed
  - Commuting distances and number of relocations in the past.
6 Future Developments

At the time this technical report was prepared, the second wave of the establishment survey was being completed and the second wave of the employee and partner surveys was in progress. Two more survey waves are planned for 2016 and 2018, provided that the SFB 882 continues to receive financial support. Use of the longitudinal design will improve the analysis of cause-and-effect relationships, and unobserved heterogeneity on the organizational and personal levels can be modeled. Additional administrative data will be updated annually.
References


IAB Integrierte Erwerbsbiographien (IEB) V11.00.00, Nuremberg 2013.


Appendix 1: List of IEB variables

The following is just a list of the LEEP-B3-data variables. A detailed description of the variables contained in the IEB extracts is provided in Berge et al. (2013).

Variables corresponding to the SIAB:

- Original start date of observation (begorig)
- Original end date of observation (endorig)
- Start date of split episode (begepi)
- End date of split episode (endepi)
- Gender (frau)
- Year of birth (gebjahr)
- Nationality, aggregated (nation_gr)
- School education and vocational training (bild)
- School-leaving qualification (schbild)
- Skills level (quali)
- Reason for notification / reason for end of benefit receipt / reason for discontinuation of unemployment benefit II / reason for deregistration (grund)
- Daily wage, daily earned rate (tentgelt)
- Transition zone (gleitz)
- Occupation – activity performed / last activity (beruf)\(^5\)
- Occupation by KldB 2010 (beruf_kldb10)\(^6\)
- Occupational status and working hours (stib)
- Employment status (erwstat)
- Employment status prior to job search (estatvor)
- Type of termination of last job (art_kuend)
- Desired working hours of the job sought (arbzeit)
- Willingness to seek work throughout Germany (mobil)\(^7\)
- Source of the observation (quelle)
- Observation counter per person (spell)

Variables that have been generated especially for LEEP B3:

- Artificial individual ID (persnr1, persnr2)
- Artificial establishment number (betnr1, betnr2)
  - Due to data protection project-specific personal and establishment numbers which do not match those in the SIAB are assigned.
    - persnr1 and betnr1 are in the IEB for establishments,
    - persnr2 and betnr2 in the IEB for persons
- Artificial partner number, link as main respondent (paarnr_mp)

\(^5\) Occupational activity according to the classification of professions 1988 (KldB 1988).
\(^6\) Occupational activity according to the classification of professions 2010 (KldB 2010). This feature is not included in the SIAB 7510.
\(^7\) This variable is included in the SIAB 7508; for a description, see Dorner et al. (2010).
- Artificial partner number, link as partner (paarnr_pa)
  o The pair numbers are used to link the IEB for persons with the employee and partner survey. Because 21 people have attended the employee as well as the partner survey, the pair number was divided in two variables.

- Main respondent or partner (mporpa)
  o indicates if a person has participated in the employee survey, at the survey, or both.
  o 1 = Main respondent
  o 2 = Partner
  o 3 = Both, main respondent and partner

- Type of link partner (linkagepa)
  o indicates the type and the quality of the link (see Section 4)
  o 1 = exact
  o 2 = exact, without street number
  o 3 = prob. match, block by postal code (5 digits), no warning
  o 4 = prob. match, block by postal code (5 digits), only with surname + forename + birth date, within one location
  o 5 = prob. match, block by postal code (5 digits), with forename + birth date + address + great surname diff.
  o 6 = prob. match, block by birth date, no warning
  o 7 = prob. match, block by birth date, only with surname + forename + birth date, within one location

- Work organization with/without establishment survey (betstat)
  o indicates if the work organization has participated in the establishment surveys
  o 1 = establishment and employer survey
  o 2 = only establishment survey

- Counter place of work: district (za Ao_kreis)
  o contains a consecutive numbering of the observed places of work of a person from 1 to the maximum number of observed places of work of this person.
  o The numbering is according to the chronological order of the first occurrence of a place of work for a person.
  o If a person returns to a previous work at a later time, this does not lead to an increase of the counter.
  o Values for older East Germans are not strictly comparable with those of older West Germans because the East German labor places are observed only from 1990 onwards.
  o The variable is only defined for employment sets.

- Counter place of residence: district (za Wo_kreis)

---

8 The information on location and district level presented in the IEB was replaced for reasons of data protection or data minimization by counters and commuting distances. The BHP extracts include the job location on the federal-state level.
- contains a consecutive number of the observed places of residence of a person from 1 to the maximum number of observed places of residence of this person.
- Residence information is only available from 1999 onwards.
- The calculation is analog to the counter work district.
- The variable is defined for all data sources.
- Distance of residence and work place (dist_wo_ao_kreis)
  - includes the distance of the district center points in km and is classified as follows: 0, >0–20, >20–40, >40–60, >60–80, >80–100, >100–150, >150–200, >200–500, >500.
  - Since information on the address are available only since 1999, the variable is only defined from 1999 onwards.
  - The variable is only defined for employment sets.

Not all variables are defined for all data sources. Full details can be found in the list of variables in Berge et al. (2013: 14). Information about workplaces and industry affiliation are included in the BHP extracts (see Appendix 2).
Appendix 2: List of BHP characteristics

The following is just a list of the LEEP-B3-data variables. A detailed description of the variables is provided in Gruhl, Schmucker, and Seth (2012).

- Core dataset
  - Establishment characteristics
    - Date of first appearance in the BeH files (grd_dat)
    - Date of last appearance in the BeH files (lzt_dat)
    - Place of work: federal state (Bundesland) (ao_bula)
    - Classification of Economic Activities 73, 2-digit code (w73_2)
    - Classification of Economic Activities 93, 2-digit code (w93_2)
    - Classification of Economic Activities 03, 2-digit code (w03_2)
    - Classification of Economic Activities 08, 2-digit code (w08_2)
  - General employment structure
    - Total number of employees (az_ges)
    - Total number of full-time employees (az_ges_vz)
    - Number of female employees (az_f)
    - Number of female employees, full-time (a_f_vz)
    - Number of German employees (az_d)
    - Number of employees by main occupation (az_hpt)
    - Number of unpaid employees (az_te0)
  - Structure of employees by educational and vocational qualifications
    - Number of low-qualified employees (az_gq)
    - Number of low-qualified full-time employees (az_gq_vz)
    - Number of qualified employees (az_mq)
    - Number of qualified full-time employees (az_mq_vz)
    - Number of highly qualified employees (az_hq)
    - Number of highly qualified full-time employees (az_hq_vz)
    - Number of employees with unknown qualifications (az_uq)
    - Number of employees with unknown qualifications (az_uq_vz)
    - Number of employees with unknown qualifications and who are not in training (az_uq_ub)
    - Number of full-time employees with unknown qualifications and who are not in training (az_uq_ub_vz)
  - Structure of employees by occupational status
    - Number of trainees/apprentices (az_az_stib)
    - Number of unskilled employees (az_nfa)
    - Number of skilled workers (az_fa)
    - Number of master craftsmen and foremen (az_mp)

9 For reasons of data protection and data minimization, the variables concerning region and economic sector in the LEEP-B3 extracts show a higher level of aggregation than do the corresponding BHP variables.
- Number of white-collar employees (az_ang)
- Number of home workers and freelance home workers (az_hh)
- Number of employees in mini part-time employment (az_ktz)
- Number of employees in midi part-time employment (az_gtz)

  Structure of employees by person group
  - Number of trainees/apprentices by person group (az_az_pers)
  - Number of marginal part-time workers (az_gf)

  Structure of employees by Blossfeld occupational group
  - Number of employees: agricultural occupations (az_bf_agr)
  - Number of employees: unskilled manual occupations (az_bf_emb)
  - Number of employees: unskilled services (az_bf_edi)
  - Number of employees: unskilled commercial and administrative occupations (az_bf_evb)
  - Number of employees: skilled manual occupations (az_bf_qmb)
  - Number of employees: skilled services (az_bf_qdi)
  - Number of employees: skilled commercial and administrative occupations (az_bf_qvb)
  - Number of employees: technicians (az_bf_tec)
  - Number of employees: semiprofessions (az_bf_semi)
  - Number of employees: engineers (az_bf_ing)
  - Number of employees: professions (az_bf_prof)
  - Number of employees: managers (az_bf_man)

  Employee age structure
  - Number of employees by age class (az_15_19, az_20_24, az_25_29, az_30_34, az_35_39, az_40_44, az_45_49, az_50_54, az_55_59, az_60_64, az_ab65)
  - Number of full-time employees by age class (az_15_19_vz, az_20_24_vz, az_25_29_vz, az_30_34_vz, az_35_39_vz, az_40_44_vz, az_45_49_vz, az_50_54_vz, az_55_59_vz, az_60_64_vz, az_ab65_vz)
  - Age quartiles of the total of employees (alter_p25, alter_med, alter_p75)
  - Age quartiles of all full-time employees (alter_p25_vz, alter_med_vz, alter_p75_vz)

  Research and development activities
  - Number of engineers and scientists (az_ingnat)

  Wage structure of full-time employees
  - Wage quartiles (gross average daily wage) for all full-time employees (te_p25, te_med, te_p75)
  - Wage quartiles for female full-time employees (te_p25_f, te_med_f, te_p75_f)
  - Wage quartiles for German full-time employees (te_p25_d, te_med_d, te_p75_d)
- Extension file – Worker flows
  - Inflows General
    - Total inflows (ein_1)
    - Total inflows: female (ein_1_f)
    - Inflows: full-time (ein_vz)
    - Inflows: full-time, female (ein_vz_f)
    - Total re-hirings (ein_wdr)
    - Re-hirings: female (ein_wdr_f)
    - Inflows: establishment movers (ein_bw)
    - Inflows: establishment movers, female (ein_bw_f)
  - Inflows by occupational status
    - Inflows: trainees / apprentices (ein_azubi_stib)
    - Inflows: unskilled workers (ein_nfa)
    - Inflows: skilled workers (ein_fa)
    - Inflows: master craftsmen, foremen (ein_mp)
    - Inflows: white-collar workers (ein_ang)
    - Inflows: mini part-time employment (ein_ktz)
    - Inflows: mini part-time employment, female (ein_ktz_f)
    - Inflows: midi part-time employment (ein_gtz)
    - Inflows: midi part-time employment, female (ein_gtz_f)
    - Inflows: marginal part-time employees (ein_gf)
  - Inflows by age class
    - Inflows: employees aged 15–19 (ein_15_19)
    - Inflows: employees aged 20–24 (ein_20_24)
    - Inflows: employees aged 25–29 (ein_25_29)
    - Inflows: employees aged 30–34 (ein_30_34)
    - Inflows: employees aged 35–39 (ein_35_39)
    - Inflows: employees aged 40–44 (ein_40_44)
    - Inflows: employees aged 45–49 (ein_45_49)
    - Inflows: employees aged 50–54 (ein_50_54)
    - Inflows: employees aged 55–59 (ein_55_59)
    - Inflows: employees aged 60–64 (ein_60_64)
    - Inflows: employees aged 65 or above (ein_ab65)
Outflows General
- Total outflows (aus_1)
- Total outflows, female (aus_1_f)
- Outflows: full-time (aus_vz)
- Outflows: full-time, female (aus_vz_f)
- Outflows: temporary (aus_temp)
- Outflows: temporary, female (aus_temp_f)
- Outflows: establishment movers (aus_bw)
- Outflows: establishment movers, female (aus_bw_f)

Outflows by occupational status
- Outflows: trainees / apprentices (aus_azubi_stib)
- Outflows: unskilled workers (aus_nfa)
- Outflows: skilled workers (aus_fa)
- Outflows: master craftsmen / foremen (aus_mp)
- Outflows: white-collar workers (aus_ang)
- Outflows: mini part-time employment (aus_ktz)
- Outflows: mini part-time employment, female (aus_ktz_f)
- Outflows: midi part-time employment (aus_gtz)
- Outflows: midi part-time employment, female (aus_gtz_f)
- Outflows: marginal part-time employees (aus_gf)

Outflows by age class
- Outflows: employees aged 15–19 (aus_15_19)
- Outflows: employees aged 20–24 (aus_20_24)
- Outflows: employees aged 25–29 (aus_25_29)
- Outflows: employees aged 30–34 (aus_30_34)
- Outflows: employees aged 35–39 (aus_35_39)
- Outflows: employees aged 40–44 (aus_40_44)
- Outflows: employees aged 45–49 (aus_45_49)
- Outflows: employees aged 50–54 (aus_50_54)
- Outflows: employees aged 55–59 (aus_55_59)
- Outflows: employees aged 60–64 (aus_60_64)
- Outflows: employees aged 65 and above (aus_ab65)

Job tenure
- Outflows: job tenure < 4 years (aus_senio_1)
- Outflows: job tenure 4–9 years (aus_senio_2)
- Outflows: job tenure > 9 years (aus_senio_3)

Extension file – entry and exit:
- Type of entry (eintritt)
  - Auxiliary variables, entry (besch, inflow, betnr_vor, besch_vor, status_vor)
- Type of exit (austritt)
  - Auxiliary variables, exit (besch, outflow, betnr_nach, besch_nach, status_nach)
Variables that have been generated specifically for LEEP-B3:
- Artificial establishment number (betnr1, betnr2)
**Additional Information on Data**

**DOI:** [10.4119/unibi/sfb882.2014.12](10.4119/unibi/sfb882.2014.12)

Creators:
- Diewald, Martin
- Schunck, Reinhard
- Abendroth, Anja-Kristin
- Melzer, Silvia Maja
- Reimann, Mareike
- Pausch, Stephanie

Title:
LEEP-B3

Publisher:
SFB 882

Year of publication:
2014

Subjects:
Sociology
DOAJ: Social Sciences
DDC: 301

Contributors:
- ContactPerson: Abendroth, Anja-Kristin
- ContactPerson: Melzer, Silvia Maja
- ProjectLeader: Diewald, Martin
- ProjectLeader: Schunck, Reinhard
- ProjectMember: Abendroth, Anja-Kristin
- ProjectMember: Melzer, Silvia Maja
- ProjectMember: Reimann, Mareike
- ProjectMember: Pausch, Stephanie
- HostingInstitution: Universität Bielefeld
- RegistrationAgency: Universitätsbibliothek Bielefeld
- Producer: Teilprojekt B3
- Funder: DFG

Dates:
- Accepted: 2014-02-13

Primary language of this resource (ISO 639):
de

Resource type:
Dataset (Stata dataset)

Size:
10000 kb
Abstract:
Previously published SFB 882 Technical Reports:


