# FDZ-DATENREPORT

Documentation of labour market data

**12|2022 EN** Establishment History Panel 1975-2021

Andreas Ganzer, Alexandra Schmucker, Jens Stegmaier, Stefanie Wolter

Bundesagentur für Arbeit

# Establishment History Panel 1975–2020

Andreas Ganzer, Alexandra Schmucker, Jens Stegmaier, Stefanie Wolter

(Institute for Employment Research (IAB))

Die FDZ-Datenreporte beschreiben die Daten des FDZ im Detail. Diese Reihe hat somit eine doppelte Funktion: zum einen stellen Nutzende fest, ob die angebotenen Daten für das Forschungsvorhaben geeignet sind, zum anderen dienen sie zur Vorbereitung der Auswertungen.

FDZ-Datenreporte (FDZ data reports) describe FDZ data in detail. As a result, this series of reports has a dual function: on the one hand, those using the reports can ascertain whether the data offered is suitable for their research task; on the other, the data can be used to prepare evaluations.

# Contents

# List of figures

# List of tables

# Zusammenfassung

Das Betriebs-Historik-Panel (BHP) setzt sich aus Querschnittsdatensätzen ab dem Jahr 1975 für Westdeutschland und ab 1992 für Ostdeutschland zusammen. Jeder Querschnitt umfasst alle Betriebe des gesamtdeutschen Raumes, die zur Jahresmitte (Stichtag: 30.6.) in der Beschäftigten-Historik (BeH) erfasst sind. Von 1975 bis 1998 sind das alle Betriebe mit mindestens einem sozialversicherungspflichtig Beschäftigten. Ab 1999 zählen zu diesen Betrieben auch solche, die zwar keine sozialversicherungspflichtigen aber zumindest einen geringfügigen Beschäftigten aufweisen. Die einzelnen Querschnittsdatensätze können zu einem Panel verbunden werden. Dieser Datenreport beschreibt das Betriebs-Historik-Panel (BHP) 1975–2021.

# Abstract

The Establishment History Panel (BHP) is composed of cross sectional datasets since 1975 for West Germany and 1992 for East Germany. Every cross section contains all the establishments in Germany which are covered by the IAB Employment History (BeH) on June 30th. These are all establishments with at least one employee liable to social security on the reference date. Establishments with no employee liable to social security but with at least one marginal part-time employee are included since 1999. The cross sections can be combined to form a panel. This data report describes the Establishment History Panel (BHP) 1975–2021.

# Keywords

German administrative micro data, labour market data, establishment data, manual

# Acknowledgements

# 1 Introduction and outline

## 1.1 Introduction

The Establishment History Panel 1975–2021 (BHP 7520) is a 50 % sample of all establishments throughout Germany with at least one employee subject to social security as of 30 June of a given year. Since 1999, establishments employing only marginal part-time workers have also been included in the data. For establishments in western Germany the observation period is 1975 to 2020 and for establishments in eastern Germany 1992 to 2021. However, conclusive analyses for eastern German establishments are only possible from 1993 onwards. The data source for the BHP is the Employee History (*Beschäftigten-Historik BeH*) of the Institute for Employment Research (IAB). The data on individuals contained in the BeH are aggregated to the establishment level using the establishment numbers. The sample of the BHP consists of between 640,000 and 1.5 million establishments per year.

The individual annual waves can be linked to form a panel dataset via the identifier "artificial establishment number", which can be found in the dataset.

This data report describes the variables of the weakly anonymous version of the BHP dataset BHP 7521, which is available to researchers via on-site use at the Research Data Centre (FDZ) or via remote data access. To a large extent the available data are the original data. In addition to the sampling procedure, only the original establishment numbers as direct identifiers of the establishments were replaced by artificial establishment IDs. In order to protect the anonymity of the establishments still further, some variables are classified as particularly sensitive and are only disclosed upon submission of a special application (see Section 1.2.1).

The data report on the BHP 7521 is structured as follows. Besides the introduction, Chapter 1 contains an outline of the data as well as information about the so-called volume structure, data use and data access. Chapter 1 also includes a short overview of the differences compared with the predecessor version BHP 7520 in order to enable users of that version to switch to the new version rapidly. A description of the data sources can be found in Chapter 2. Data preparation and data quality are discussed in Chapters 3 and 4, whilst the individual variables are described in detail in Chapter 5. The appendix (Chapter 8) provides detailed descriptions of data preparation procedures and information on further working tools.

## 1.2 Data access and use

### 1.2.1 Data access

The BHP data in the weakly anonymous version may be analysed in the context of a research visit at the Research Data Centre of the BA at the IAB (FDZ) and subsequent remote data access. It is also possible to analyse the data solely via remote data access.

In order to be able to use the data, in either case it is first necessary to submit an application to the FDZ. Details on how to apply for data access can be found on the FDZ website (https://fdz.iab.de).

In addition to the core dataset, which contains the stock information for the years from 1975 to 2021, there are also the extension files "worker flows", "entry and exit" and "bankruptcies". All of the extension files require additional applications for data access.

Certain variables which make it possible to identify establishments are only disclosed in their original form if this is necessary for the analysis objective and is justified explicitly in the application for data access. The variables in the core dataset which are particularly sensitive from the viewpoint of data protection legislation are

- place of work: district (ao_kreis)

- economic activity 93 - sub-class of economic activity (w93_5)

- economic activity 03 - sub-class of economic activity (w03_5)

- economic activity 08 - sub-class of economic activity (w08_5)

Details regarding the place of work and the economic activity are available in aggregated form (federal state and economic activity as 3-digit codes) in the core dataset without special application, however.

In addition, establishment-specific effects (AKM effects) are available for the BHP 7521 for the years 1985 to 2021. Further information can be found in the FDZ Methodenreport 01/2023 (Lochner et al., 2023).

## 1.2.2  Data use

The BHP is available as a 50 % simple random sample. The BHP data are divided into five modules. The core dataset comprises simple establishment variables such as the artificial establishment ID, the federal state or the economic sector in which the establishment is active as well as details about the employee structure and the wage structure of full-time employees. The core dataset is available as yearly data for the years 1975 to 2021. The yearly data can be merged to form a panel dataset via the artificial establishment ID. The files of the core dataset can also be supplemented by the variables for the time-consistent classifications of economic activities, which are stored in a separate file. In addition to the core dataset there are three extension files, which can be merged with the core dataset via the artificial establishment ID. The extension file "worker flows" contains details on the annual inflows and outflows of workers. The second extension module "entry and exit" contains detailed information about the entry of the establishment and its exit if applicable. The third extension module "bankruptcies" contains information that allows identifying establishments that vanish from the data due to bankruptcies. The following programming example shows how a panel dataset can be created from the waves of the core dataset and supplemented by additional information from the time-consistent classifications of economic activities and the extension files.

```
* Programming example for creating a panel dataset:
* Merging the waves of 2003-2021
* Generating the variable jahr in order to distinguish between waves and to merge extension
files
* Selecting the required variables bhp_betnr, az_ges and az_f
clear
foreach y of numlist 2003/2021 {
        append using $orig/bhp_7521_m06_p50_v1_`y'.dta, keep(bhp_betnr az_ges az_f)
        cap gen jahr=.
        replace jahr=`y' if jahr==.
```

```
}
* Generate year of last occurrence in BHP (only to merge bankruptcies)
bysort bhp_betnr (jahr): gen lzt_jahr=jahr[_N]

* Merging the time-consistent classification of economic activities WZ93
merge 1:1 bhp_betnr jahr using $orig/bhp_7521_m06_p50_wgen_v1.dta, keepusing(w93_3_gen)
drop if _merge == 2
drop _merge

* Merging the inflows and outflows
merge 1:1 bhp_betnr jahr using $orig/bhp_7521_m06_p50_inflow_v1.dta, keepusing(ein_ges
ein_f)
replace ein_ges = 0 if missing(ein_ges)
replace ein_f = 0 if missing(ein_f)
drop if _merge == 2
drop _merge
merge 1:1 bhp_betnr jahr using $orig/bhp_7521_m06_p50_outflow_v1.dta, keepusing(aus_ges
aus_f)
replace aus_ges = 0 if missing(aus_ges)
replace aus_f = 0 if missing(aus_f)
drop if _merge == 2
drop _merge

* Merging the entries and exits
merge 1:1 bhp_betnr jahr using $orig/bhp_7521_m06_p50_entry_v1.dta, keepusing(eintritt)
drop if _merge == 2
drop _merge
merge 1:1 bhp_betnr jahr using $orig/bhp_7521_m06_p50_exit_v1.dta, keepusing(austritt)
drop if _merge == 2
drop _merge

* Merging the bankruptcies
merge m:1 bhp_betnr using $orig/bhp_7521_m06_p50_iv_v2.dta
* keep bankruptcy information for last year only
foreach v of varlist iv_monitoring-iv_geld_endjahr {
        replace `v'=. if lzt_jahr!=jahr
        }
drop if _merge == 2
drop _merge lzt_jahr

compress
save $data/bhp_panel_0321.dta, replace
```

The BHP data contain both German and English labels. Using the Stata command "`label language en`" or "`label language de`" the labels can be switched to English or German.

In order to minimise the memory requirements of the BHP, some variables are not explicitly included in the data if they can be calculated by the users themselves using the available information (see Section 3.10).

## 1.3 Differences compared with the predecessor version

### 1.3.1 Data basis

**Update**

For the update of the BHP 7521, the data for all of the years recorded were newly generated, i.e. also the existing waves of the old BHP from 1975 to 2020 were compiled again. The source of the data for the BHP, the BeH, is regularly updated in its entirety, which is why discrepancies may occur

between the different BHP versions in the waves of 1975–2020. The 2021 wave is a special case in this respect as information is incorporated into the BHP after a waiting period of only 6 months here and experience has shown that the notifications cannot be assumed to be complete (see Section 4.3).

**Territorial allocation**

The regional information is now given with the territorial allocation as of 31.12.2021.

**Deleted and new variables**

The variables „bhp_betnr_vor" (Establishment id predecessor) and „bhp_betnr_nach" (Establishment id successor) are no longer provided.

**Changed generation of the variables on censored and imputed wages**

A new procedure is used to determine the employment episodes with censored wage information. The upper earnings limits are determined for each year (separately for East and West Germany) as 97.5% of the mode of daily wages of full-time employees rounded to whole Euro amounts. This has an effect on the variable "az_zens" (Number of employees with censored wages) and on all variables containing information on imputed wages ("te_imp_").

### 1.3.2 Data Report

As part of the revision of the data report, the changes described above were included. The sections for the removed variables were deleted and a new section was created for the new variable.

## 1.4 Outline

**Table 1: Outline**

| Categories | Descriptions |
|---|---|
| **Topics** | Core dataset:<br>• establishment characteristics (e.g. artificial establishment number,<br>• economic activity, federal state)<br>• structure of employees (e.g. by gender, form of employment)<br>• employee age structure<br>• structure of employees by educational and vocational qualifications<br>• structure of employees by occupational group (Blossfeld classification of occupations)<br>• average wage of full-time employees<br>• sum of wages of all employees<br>Extension file – worker flows:<br>• total number of inflows<br>• structure of inflows by gender, form of employment, Blossfeld classification of occupations, number of re-hires, age<br>• total number of outflows<br>• structure of outflows by gender, form of employment, Blossfeld classification of occupations, job tenure, age<br>Extension file – entry and exit<br>• entry year<br>• type of entry<br>• auxiliary variables for defining entry type |

| Categories | Descriptions |
|---|---|
| | • exit year<br>• type of exit<br>• auxiliary variables for defining exit type<br><u>Extension file – bankruptcies (from 2007 onwards)</u><br>• sources of bankruptcies information<br>• begin and end date of the bankruptcies information |
| **Research unit** | Establishments in Germany with at least one employee subject to social security or (since 1999) with at least one marginal part-time employee on 30 June |
| **Number of cases** | Annual number of observations: 640,000 – 1.5 million establishments |
| **Period covered** | Western Germany: 1975 – 2021<br>Eastern Germany: 1992 – 2021 |
| **Time reference** | 30 June of each year |
| **Regional structure** | Federal states (Bundeslaender), districts (Kreise) |
| **Type of territorial allocation** | Corrected territorial allocation as of 31/12/2021 |
| **Data collection method** | 50 % random sample |
| **Institutions involved** | Social security agencies, Federal Employment Agency |
| **Frequency of data collection** | Continuous |
| **File format and size** | Stata<br>Core dataset: each annual wave 190–470 MB<br>Time-consistent Classification of Economic Activities: 740 MB<br>Extension file – worker flows: 2 x 2.5 GB<br>Extension file – entry and exit: 2 x 55 MB<br>Extension file – bankruptcies:2.8 MB |
| **File architecture** | Core dataset: annual waves<br>Time-consistent Classification of Economic Activities: 1 file<br>Extension file – worker flows: 2 files (inflows, outflows)<br>Extension file – entry and exit: 2 files (entries, exits)<br>Extension file – bankruptcies: 1 file |
| **Data access** | On-site use or remote data access |
| **Anonymisation degree** | Weakly anonymous |
| **Sensitive variables** | District (Kreis), classification of industries WZ93, 5-digit; classification of industries WZ03, 5-digit; classification of industries WZ08, 5-digit |
| **Standards for the citation of data and data documentation** | **Data**:<br>Ganzer, Andreas; Schmucker, Alexandra; Stegmaier, Jens; Wolter, Stefanie (2022): „Establishment History Panel 1975–2021 (BHP 7521 v1)". Research Data Centre of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB). DOI: 10.5164/IAB.BHP7521.de.en.v1<br>The data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and/ or remote data access.<br>**Data documentation:**<br> - Andreas Ganzer;  Schmucker, Alexandra, Stegmaier; Jens; Wolter, Stefanie (2022): Betriebs-Historik-Panel 1975–2021. FDZ-Datenreport, 12/2022 (de), Nürnberg. DOI: 10.5164/IAB.FDZD.2212.de.v1<br> - Andreas Ganzer;  Schmucker, Alexandra;  Stegmaier, Jens; Wolter, Stefanie (2022): Establishment History Panel 1975-2021. FDZ-Datenreport, 12/2022 (en), Nürnberg. DOI: 10.5164/IAB.FDZD.2212.en.v1 |

| Categories | Descriptions |
|---|---|
| **Current data version** | The Establishment History Panel 1975–2021 (BHP 1975-2021); DOI: 10.5164/IAB.BHP7521.de.en.v1 |

The dataset described in this document is available for use by professional researchers. Further information can be found on the website https://fdz.iab.de/.

## 1.5   Volume structure

**Table 2: Volume structure**

50% simple random sample

| File name | File size | Number of observations | Number of establishments |
|---|---|---|---|
| **Core dataset** | | | |
| bhp_7521_m06_p50_v1_1975 | 189 MB | 643,761 | 643,761 |
| ... | … | … | … |
| bhp_7521_m06_p50_v1_2021 | 454 MB | 1,482,144 | 1,482,144 |
| **Time-consistent Class. of Economic Activities** | | | |
| bhp _7521_m06_p50_wgen _v1 | 738 MB | 51,641,266 | 5,525,188 |
| **Extension file - worker flows** | | | |
| bhp_7521_m06_p50_inflow_v1 | 2.5 GB | 30,292,424 | 5,525,188 |
| bhp_7521_m06_p50_outflow_v1 | 2.3 GB | 28,563,785 | 5,195,255 |
| **Extension file - entry and exit** | | | |
| bhp_7521_m06_p50_entry_v1 | 54 MB | 3,572,652 | 3,572,652 |
| bhp_7521_m06_p50_exit_v1 | 50 MB | 3,301,737 | 3,301,737 |
| **Extension file - bankruptcies** | | | |
| bhp_7521_m06_p50_iv_v1 | 2.8 MB | 168,191 | 168,191 |

# 2 Data sources

## 2.1 Employee History (BeH)

The source of data regarding employment is the Employee History (*Beschäftigten-Historik* – BeH) of the IAB. The data basis is the integrated notification procedure for health, pension and unemployment insurance, which came into effect as of 1 January 1973 (and was extended to cover eastern Germany as of 1 January 1991) and is known by the abbreviation DEÜV (previously DEVO/DÜVO) (for further details see: Bender et al. 1996, p. 4 ff.; Wermter and Cramer 1988). Under this procedure employers are required to submit notifications to the responsible social security agencies concerning all of their employees covered by social security. The BeH covers all white- and blue-collar workers as well as apprentices as long as they are not exempt from social security contributions. This means that civil servants, the self-employed and regular students (see Cramer 1985) are in principle not recorded in the BeH. Since the notification procedure was changed on 1 January 1999, employees in marginal part-time employment and unpaid family workers have also been recorded (not contained in the data until 1 April 1999).

Every year in which an individual is in an employment relationship is depicted by at least one notification. The data are recorded by the health insurance companies, collected in a continuous file by the Federal Employment Agency (BA) and subsequently integrated into the History File of the IAB. The current observation period of the BeH extends from 1 January 1975 to 31 December 2021.

## 2.2 Establishment and establishment number

The information about establishments contained in the administrative data of the Federal Employment Agency is based on a specific definition of 'establishment'. According to this definition, an establishment is a regionally and economically delimited unit in which employees work. An establishment may consist of one or more branch offices or workplaces belonging to one company. The term 'company' combines all establishment premises and workplaces belonging to the same employer. An 'employer' is any natural person or legal entity that is the party liable for the overall social security contribution and employs at least one employee subject to social security contributions or in marginal part-time employment (see Bundesagentur für Arbeit 2007).

In order to identify establishments, they are each allocated a unique establishment number by the Federal Employment Agency.[1] The establishment number is generated automatically by computer and consists of eight digits. The first seven digits of the establishment number constitute a serial number; the eighth digit is a check digit. This original establishment number is not used in the microdata of the Research Data Centre. Instead, an artificial establishment ID is used which does not permit any re-identification of the original number.

Only establishments that are obliged to submit social security notifications are given an establishment number. The following principle applies for the allocation of establishment numbers: branch offices of one company which belong to the same economic class and are located

---

[1] Since 1.10.2007 the establishment numbers have been allocated by the central establishment number service of the BA, prior to that date the local employment agencies were responsible.

in the same municipality are given one joint establishment number. It is not possible to distinguish between branch offices with a joint establishment number in the data. Furthermore, no information is available as to whether establishments belong to the same company.

When the establishment number is assigned, the establishment is also classified according to the Classification of economic activities of the German Federal Statistical Office. The classification depends on the purpose of the establishment and the main economic activity of the establishment, i.e. there is always only one unambiguous assignment per establishment to an industry. The allocation is based on the information provided by the employer. Changes in the classification of economic activity are possible and are usually reported to the BA by the establishment themselves.

Once an establishment has been allocated an establishment number it remains unchanged in principle. This also applies if the establishment relocates to another municipality. Occasionally an existing establishment is given a new establishment number if …

- there is a change of owner (not, however, if the establishment is transferred to another member of the family),

- the company is split up or

- the company merges with another company.

A new establishment number is not allocated, however, if a company is taken over by another company.

In addition, a new establishment number can be allocated on application if …

- there has been a change in the legal form,

- the establishment is transferred to another member of the family,

- insolvency proceedings have been initiated or

- a "dormant" establishment is being reactivated[2].

If the economic class of a branch office changes, the case is examined to check whether a new establishment number is to be allocated. This is not necessarily so.

In practice, however, it is possible to diverge from the principles of allocating and changing establishment numbers, provided that the employer and/or the health insurance company agree.

Using the establishment number, the particular establishment can be matched with its location, its economic activity and its reported employees subject to social security and in marginal employment.

The first and last appearance of an establishment number offer a first indication for the times when the establishment was founded and closed down. However, the establishment number carries no information on changes in the structure of the branch offices, establishments and companies (splits, fusions, restructuring, etc.). By using worker flows between establishment numbers it is possible to distinguish between genuine start-ups and closures and changes in establishment numbers, takeovers or spin-offs (see Hethey and Schmieder 2010, 2013). The corresponding

---

[2] A 'dormant' establishment means either that there are no employees working there for whom social security notifications have to be submitted or that the establishment has closed down. The establishment can be reactivated either by hiring new employees or by being founded again by the previous owner.

variables for this classification can be found in the "entry and exit" extension file (see Section 3.5 and Section 5.3).

# 3 Data preparation

## 3.1 Corrections and validation procedures performed on data on individuals

For compiling the BHP, the employment notifications of the BeH (see Section 2.1) are aggregated at establishment level using the establishment ID (see Section 2.2). Before this aggregation, the data on individuals are subjected to numerous validation procedures. These are described in detail below.

### 3.1.1 Selection of the notifications in the BeH

The data on individuals from the BeH are used as the basis for the BHP, but not all the notifications are included:

- Only notifications with details about the following person groups are taken into account: 101,102, 103, 105, 106, 109, 112, 113, 114, 118, 119, 120, 121, 122, 140, 141, 142, 143, 144, 149, 150, 201, 203, 205, 209, 999, YYY (see Section 8.2).

- Notifications with a wage of 0 are deleted. As these notifications concern deregistrations of individuals who were previously sick or on parental leave and received corresponding earnings replacement benefits, these individuals are not counted as employees.

- Notifications before 1992 reporting a place of work in an eastern German federal state (excluding Berlin) are deleted, as the social security notifications for eastern Germany can only be assumed to be complete from 1993 onwards (see Section 4.1). Data for 1992 are included but not complete.

### 3.1.2 Strike corrections

In the spring of 1984 there were strike-related lockouts in establishments in the "manufacture of motor vehicles, motor vehicle engines" (WZ73: 280) and "manufacture of parts and accessories for motor vehicles" (WZ73: 281) industries in Hesse and Baden-Wuerttemberg, which is reflected in the data on individuals in the form of gaps in employment. As these gaps frequently also include the reference date of 30.06.1984, which is relevant for the BHP, this would have resulted in considerable distortions in the BHP for the industries in the federal states affected in 1984. These gaps were therefore filled in accordance with the following heuristics:

First, the gaps resulting from lockouts had to be identified. The following definition was used for this.

An account was regarded as locked out if ...

- there was a notification in Baden-Wuerttemberg or Hesse on 30.04.1984 that was classified as belonging to the economic activity 280 or 281 (notification 1),

- there was in addition a further notification from the same establishment in July 1984 (notification 2) and

- in May or June 1984 there was a gap in employment lasting more than 5 days.

These gaps were filled by transferring the start date in notification 1 to notification 2 and adding together the earnings details from the two notifications. Then notification 1 was deleted. If there were further notifications between the first and the second notifications, they were also deleted and the earnings details were added accordingly.

### 3.1.3    Validation of the information on earnings

#### 3.1.3.1    Addition of special payments

As a rule, the employers already include any special payments (such as holiday pay, 13th monthly salary etc.) in their regular annual notifications or deregistrations. In some cases, however, the special payment is reported separately (notification reason 54). These payments, too, have to be taken into account when calculating the earnings data of an establishment; for this, the earnings of the extra notification are added to the earnings of the regular notification in the same calendar year. If there are no such regular notifications, the special payment is disregarded when compiling the BHP.

#### 3.1.3.2    Completing missing information on earnings

In the period 1992–1998 notifications without earnings details can be found in the mining sector. As the other variables in these notifications contain valid information, it can be assumed that the jobs did actually exist. Perhaps problems occurred when the earnings were reported. In order to fill in the missing earnings information, the following procedure is implemented:

- Continuation: if the episode concerned is preceded by a period of employment in the same establishment with an annual notification (reason for notification = 50) and with the same person group, and there is no gap between these two episodes (i.e. a gap of 0 days), then the earnings from the preceding episode are carried forward. If there are several consecutive episodes without information on earnings and if the conditions described above are also met, the last available earnings are carried forward in each case. In this way 95% of the missing values can be filled in.

- Writing back: for the episodes that still have missing information on earnings after the continuation procedure, the earnings from the following observation are carried back. The condition for this is that the episode concerned is followed by a period of employment in the same establishment with an annual notification (reason for notification = 50) and with the same person group, and that there is no gap between these two episodes (i.e. a gap of 0 days). In this way the remaining 5% of the missing values can be filled in.

#### 3.1.3.3    Imputation of data on earnings above the upper earnings limit

In the social security notifications, earnings are only reported up to the upper earnings limit for statutory pension insurance contributions. This means that approx. 10 % of the information on full-time employees' earnings is censored[3]. This leads to biased results due to aggregation: means of earnings are biased if the censored observations are not included in the calculation or if censored values are replaced by the censoring limit. No bias occurs for wage quantiles below the censoring limit. As the shares of censored wages can sometimes be very large (well over 50%) depending on

---

[3] The upper earnings limits are determined for each year (separately for East and West Germany) as 97.5% of the mode of daily wages of full-time employees rounded to whole Euro amounts.

the wage level in the establishment, in many analyses it would only be possible to use quantiles below the median. In order to remedy these two problems, the information on earnings (average daily wages) were imputed before the statistics (means and medians) were calculated. The procedure is implemented following Card et al. (2015) and is explained in more detail in Section 8.1 in the Appendix.

### 3.1.4   Validation of the data on education and vocational training

The number of employment notifications with missing information on education and vocational training qualifications has grown substantially over time; this concerns people in marginal part-time employment to a disproportionately large extent. The switch to the Occupation Code 2010 (see Section 4.6) in the notification procedure caused the rate of missing values to rise as high as 50% in 2011. Furthermore, from 2011 onwards the employers no longer report qualifications in a combined variable, but split into school education (none, lower secondary, intermediate secondary, upper secondary) and vocational education and training (none, recognised vocational training, master craftsman, bachelor, diploma, doctorate). This actually permits a more precise recording of education and training qualifications, but no time-consistent information is available for the entire period. In order to achieve that, the methods of recording the data have to be made compatible. This is done by assigning to every combination of values from the new occupation code the most suitable details on education and training according to the old occupation code (see Thomsen et al., 2018). This has no effect on missing values, however.

In addition, the evaluability of the education and training data is improved by means of an imputation procedure using a deterministic replacement rule that was suggested by Fitzenberger et al. (2005, 2006) and enhanced by Thomsen et al. (2018). The result of this procedure is that there are now hardly any missing values, especially for employees who are not in marginally part-time employment. A precise description of the initial situation, the various imputation procedures that were tested, and the final outcome can be found in Thomsen et al. (2018).

### 3.1.5   Validation of the occupation data

In order to obtain time-consistent information on occupations for periods of time with an end date after 30.11.2011, the 5-digit occupational class according to the Classification of Occupations 2010 (KldB 2010) was coded back to the structure of values found in the Classification of Occupations 1988 (KldB 1988). As this is only possible on the basis of modes, inaccuracies occur when assigning occupations to classes. The quality problems detected can basically be put down to two causes. First, updating effects emerge, as the introduction of the new occupation code prompts employers to check and, if necessary, to update the information they have so far reported about their employees. Second, the transcoding from the new system to the old one produces inaccuracies. For instance, several different occupational groups (3-digit according to KldB 1988) can be assigned to one occupational class (5-digit according to KldB 2010) (see Paulus and Matthes 2013a, 2013b).

On the basis of these recoded occupation variables it was possible to form the 12 categories of the Blossfeld classification of occupations (see Blossfeld 1987) for the years from 2011 onwards, too. The variable "az_ingnat" (number of engineers and natural scientists) is generated on the same basis.

Furthermore, in the reporting year of 2011 some 8 million notifications (approx. 15 %) contain no valid information on the occupation. In the reporting year of 2012 about 1.6 million notifications, i.e. approx. 3 %, are affected by the problem described above. Unlike with the information on working hours and on education and vocational training qualifications, no method has yet been developed for imputing the missing values for occupations, so especially in 2011 the information in the variables concerning occupations according to the Blossfeld classification are incomplete (see Section 4.6).

As the new classification of occupations (KldB 2010) also records information about the level of requirement in the last position, which was not depicted in the previous classification of occupations (KldB 1988), this information is to be used for the BHP. The variables "az_niv1", "az_niv2", "az_niv3" and "az_niv4" show the number of employees with different levels of requirement. Due to the known missing values, the information for 2011 is incomplete (see Section 4.6).

### 3.1.6 Validation of the information on full-time and part-time employment

As a result of the introduction of the new occupation code on 01.12.2011 the data no longer distinguish between "midi" and "mini" part-time employment. In order to obtain time-consistent information, now only the total number of part-time employees is shown, also for years before 2011. Unlike the old versions of the BHP, the versions from BHP 7514 onwards do not use employees in marginal part-time employment, apprentices, or people participating in partial retirement schemes to generate the full-time and part-time variables (see Sections 5.1.2.7 and 5.1.2.8).

Moreover, for a transitional period after the introduction of the new occupation code it was permitted to leave out the information on the occupation code in the social security notifications (see Section 4.6). In a good 10 % of the notifications submitted by the establishments between 01.12.2011 and 31.05.2012 the information regarding working hours is therefore missing. For this reason, a method l was developed at the IAB which can be used to impute the missing information. The imputation model for the part-time variable was estimated by using a recursive binary splitting algorithm which created a binary classification tree (see Ludsteck and Thomsen 2016). The information on working hours that is generated using this procedure is contained in this variable.

### 3.1.7 Correction of misclassifications in the municipality of Lahn

There were far too many employees with a place of work in the Lower Saxony municipality of Lahn (district of Emsland) in the years 1975-1977. Instead of the expected two-digit figure, there were up to 90,000 employee notifications per year with this place of work. The reason for this is a historical misclassification of employment notifications from the city of Lahn. This was an incorporation of the Hessian municipalities Wetzlar, Gießen, Heuchelheim, Wettenberg and Lahnau (districts of Gießen and Lahn-Dill-Kreis) which was dissolved after a short time.

Based on these findings and assuming that there were no real establishment relocations between these regions, the following adjustment rule was implemented on the level of the establishment number. As soon as the establishment location changes from the municipality of Lahn to one of the listed Hessian municipalities in the years 1975 to 1978, the former information on the

municipality of Lahn is replaced with the more recent, corrected information. This rule significantly reduces the surplus of employment notifications and the municipality of Lahn in Lower Saxony then has at least less than 2,000 employees in the concerning years. The misclassification was corrected for BHP 7517 for the first time. A detailed documentation on how numbers in the affected districts changes is available in Schmucker et al. (2018).

## 3.2 Generating the core dataset

The core dataset contains information on the establishment's employee structure and details about the wage structure of the full-time employees in the establishment. All of the data are stock data as of the reference date of 30.06 of any given year. After the data on individuals have been preprocessed (see Section 3.1) the core dataset is generated as follows:

- Selection of all BeH observations that include the reference date of 30.06 of the respective year.

- Conversion of all earnings into euros.

- Deletion of multiple jobs held by one person in one and the same establishment: Here non-marginal jobs are given priority over marginal part-time jobs. If more than one non-marginal or marginal job is recorded for one person in the same establishment, the job with the higher wage is selected or, if the jobs have the same wage, a random selection is made.

- Aggregation of all employment notifications as of the reference date of 30.06 of any year to form selected statistics at establishment level on the basis of the establishment ID. Due to this procedure based on reference dates it may happen, for example, that establishments which closed down before the reference date do not appear in the corresponding block of yearly data in the BHP although valid employment notifications are still available in the BeH before 30.06. The same applies for establishments that were not founded until the reference date. These might not appear in the BHP until a year later.

- Storage of the data in separate files for each year. All of the variables are contained in all the annual files, even if they only show missing values (e.g. number of employees in marginal part-time employment before 1999).

After the aggregation process variables and values are given German and English labels and less detailed variables are generated from existing variables (federal state, 3-digit code for the economic activity).

## 3.3 Generating the time-consistent classifications of economic activities

The Classification of Economic Activities has been modified several times during the observation period of the BHP, which renders longitudinal analyses difficult. The FDZ developed a method to generate time-consistent industry codes. A detailed description of generating time-consistent codes for the classification of economic activities WZ73 and WZ93 can be found in Eberle et al. 2011. Since version BHP 7510, time-consistent industry codes for the classifications WZ73 and WZ93 have also been compiled for periods in which industry information is given only for the classification Classification of Economic Activities WZ08.

For the present version of the BHP, the established variables have been updated. In addition, a time-consistent variable has been computed for the Classification of Economic Activities WZ08 by extrapolation and imputation. The method used to generate the time-consistent codes corresponds to the procedures described by Eberle et al. 2011 for the generation of time-consistent information for WZ73 and WZ93. The generated industry codes are available in an additional dataset and can be merged via the artificial establishment ID and the year (see Section 1.5).

In order to enter industry codes according to WZ08 for observations prior to 2008, values are first extrapolated within an establishment over time wherever possible and subsequently imputed based on the computed correspondence tables. This procedure is not exact and there may be misclassifications. The propensity of misclassifications is especially high for earlier years, since fewer values can be extrapolated and values have to be transcoded across multiple classification schemes.

## 3.4 Generating the extension file – worker flows

In addition to details about the number of employees in the establishments as of the reference date of 30.06 of any given year, the BHP also contains details of worker flows. The data distinguish between employee inflows and outflows. The inflows of a year are defined as the number of employees who were working in the establishment on the reference date of that year but were not working there on the reference date of the previous year. Analogously, the outflows of a year are defined as the number of employees who were not working in the establishment on the reference date of the particular year but were working there on the reference date of the previous year. Thus, the current number of employees in a particular year, as it is found in the core dataset, can be calculated as the number of employees in the previous year plus all inflows minus all outflows of the current year.

Example: an establishment has 30 employees as of 30.06.1995. 5 employees leave the establishment between 30.06.1995 and 30.06.1996. These 5 employees are therefore the outflows for 1996. At the same time the establishment hires 7 new employees. These are the inflows for 1996. The total number of employees for 1996 is therefore 30 - 5 +7 = 32.

Employees who join the establishment and leave it again between two reference dates are not recorded by this flow concept based on reference dates. Changes in a person's status within an establishment are not shown either, e.g. when a person is reported as employed in the same establishment on two consecutive reference dates but with different occupations.

Similar to the stock values in the core dataset, the inflows and outflows are broken down according to various characteristics such as age groups. Further information on the individual variables can be found in Chapter 5.2.

Unlike in the core dataset, the worker flows are not filed as blocks of yearly data but in two files (one for entries, one for exits), which can be merged with the individual sets of yearly data of the core dataset via the establishment number and the year. When merging this data, the following restrictions have to be taken into account:

- Establishments with no inflows or outflows in a year only appear in the core dataset.

- In an establishment's year of entry, the inflows are equivalent to the employee stocks in the core dataset.

- Establishments that have closed down appear only in the worker outflow dataset in the following year. The outflows listed here are equivalent to the employee stocks of the preceding year (year of exit).

The variables of worker flows contain hardly any missing values. Missing values appear only if details are systematically not available. For example, marginal part-time jobs were not recorded before 1999. Thus, worker flows cannot be calculated before 1999. Analogously, inflows or temporary outflows cannot be calculated with the last available data due to missing values of the next year.

Besides, it must be taken into account that the value 0 could have two different meanings. On the one hand the value 0 is assigned if there are no inflows or outflows between the two reference dates for the respective person group. On the other hand, the value 0 can also be attributed to an employee group not represented in the establishment. To distinguish between the two meanings, the worker flows could be merged with the core dataset[4].

The extension file 'worker flows' is only made available to users on application.

## 3.5 Generating the extension file – entry and exit

This dataset includes information about the classification of establishment start-ups and establishment closures. This information should help to identify the start-up of a completely new establishment or the closure of an establishment in contrast to parts of an establishment being split off or outsourced, or an existing establishment simply being given a new ID.

### 3.5.1 Generation and content

Using the variables in the core dataset of the BHP it is very difficult to identify or classify establishment start-ups and establishment closures. The first and last appearance of an establishment number as they are shown in the core dataset were used as proxies for the start-up and closure dates. As these two points in time only indicate the period during which the establishment reported employees covered by social security or marginal part-time employees they are not necessarily the dates when the establishment was set up or closed. Apart from the start-up of a completely new establishment, the appearance of a new establishment number in the dataset can also mean that parts of an establishment have been split off or outsourced, or simply that an existing establishment has been given a new establishment number (see Section 2.2). (Regarding the allocation and changes of establishment numbers see Bundesagentur für Arbeit (2007: 9–11).)

---

[4] E.g.: To distinguish whether an establishment had no outflow of marginal part-time employees from 2009 to 2010 or there were no marginal part-time jobs in the establishment in 2009, one could recode aus_gf in Stata with the following programming:

```
use bhp_betnr az_gf using bhp_7521_m06_p50_v1_2009.dta, clear
save temp.dta, replace
use bhp_betnr jahr aus_gf using bhp_7521_m06_p50_outflow_v1.dta, clear
sort bhp_betnr
merge 1:1 bhp_betnr using temp.dta
replace aus_gf = . if az_gf == 0
```

Within a cooperation project with the University of Boston an attempt has been made to classify establishment entries and exits more precisely on the basis of worker flows, thereby distinguishing between genuine start-ups and closures and cases in which jobs are simply shifted. The decisive factor for the classification of establishment entries is the proportion of workers who were employed in one and the same establishment in the previous year (maximum clustered inflow). If this proportion is very large and the predecessor establishment no longer exists, it can be assumed that the new establishment is merely an existing establishment that has been allocated a new establishment number. If, on the other hand, the workers are recruited from a multitude of different establishments, there is a greater likelihood that the case is a genuine new establishment. Analogously, it applies for establishment exits that if the proportion of workers that jointly switched to a new establishment in the following year increases (maximum clustered outflow), then the likelihood of this also simply being a case of a new establishment number and not a closure also increases.

The data basis used for this classification is an additional dataset containing all the worker flows occurring between two consecutive reference dates for the years from 1975 to 2020. Unlike the data basis of the BHP, only employees covered by social security and no marginal part-time employees are taken into account for this in order to obtain a time-consistent basis. Furthermore, only one job per reference date, the so-called main employment, is included for each person.

To classify newly founded establishments, a distinction is first made with regard to how many employees the new establishment has in the year of its first appearance. As it makes little sense to examine the worker flows of establishments with fewer than 4 employees, these are generally classified as newly founded establishments. In the case of larger establishments, it is then examined whether the establishment from which the largest cluster of workers switched to the new establishment continues to exist or not. Finally, a further classification is made using the maximum clustered inflow as a share of the total workforce of the new establishment (see Table 2). The classification of establishment closures is done along the same lines. Here, too, the disappearance of an establishment with fewer than 4 employees is always defined as a "genuine" closure. Larger establishments are classified separately depending on whether the successor establishment already existed previously or not (see Table 3).

Using this information, it is possible to classify establishment entries in western Germany for the period 1976-2021 and establishment entries in eastern Germany for 1993-2021. Establishment exits in western Germany can be classified for the period 1975-2020 and those in eastern Germany for 1992-2020.

A detailed description of the classification can be found in Hethey and Schmieder (2010), which can be downloaded free-of-charge from the FDZ homepage. All users who wish to work with the extension file 'entry and exit' are strongly recommended to read this Methodenreport.

The extension file contains all of the variables necessary for the classification, so the user can either adopt the developed classification in its original form or can modify it in certain places and adapt it to his own interests. It is possible, for example, to calculate the share of the maximum clustered inflow as a percentage of the workforce of the new establishment using the variables "besch" and "inflow". Using this information, it is then possible, for instance, to implement other definitions of changes in establishment numbers and spin-offs, and to change the "eintritt" variable. The information on the establishment history is stored in two files, which can be merged

with the yearly data in the core dataset via the establishment ID and the year in which the establishment ID appeared for the first and/or the last time.

**Table 3: Classifying entering and exiting establishments**

| Entries | MCI / Inflows | Predecessor exits — MCI / Predecessor Employment | | | Predecessor continues — MCI / Predecessor Employment | | | No predecessor MCI=0 |
|---|---|---|---|---|---|---|---|---|
| | | < 30 % | 30-80 % | > 80 % | < 30 % | 30-80 % | > 80 % | |
| ≤ 3 empl. | - | New (small) | New (small) | New (small) | New (small) | New (small) | New (small) | New (small) |
| > 3 empl. | < 30 % | New (mid / big) | New (mid / big) | New (mid / big) | New (mid / big) | New (mid / big) | New (mid / big) | New (mid / big) |
| | 30-80 % | New (chunky) | New (chunky) | Unclear | New (chunky) | New (chunky) | Unclear | |
| | > 80 % | Spin-off pushed | Spin-off pushed | ID Change | Spin-off pulled | Spin-off pulled | Unclear | |

MCI = Maximum Clustered Inflow

| Exits | MCI / Outflows | Successor is entrant — MCO / Successor Employment | | | Successor is existing estab. — MCO / Successor Employment | | | No successor MCO=0 |
|---|---|---|---|---|---|---|---|---|
| | | < 30 % | 30-80 % | > 80 % | < 30 % | 30-80 % | > 80 % | |
| ≤ 3 empl. | - | Death (small) | Death (small) | Death (small) | Death (small) | Death (small) | Death (small) | Death (small) |
| > 3 empl. | < 30 % | Death (Atomized) | Death (Atomized) | Spin-off pushed | Death (Atomized) | Death (Atomized) | Death (Atomized) | Death (Atomized) |
| | 30-80 % | Chunky Death | Chunky Death | Spin-off pushed | Chunky Death | Chunky Death | Chunky Death | |
| | > 80 % | Unclear | Unclear | ID Change | Take-Over / Restruct. | Take-Over / Restruct. | Unclear | |

MCO = Maximum Clustered Outflow

See Hethey and Schmieder (2010), p. 24

### 3.5.2 Differences to the characteristics "grd_dat" and "lzt_dat"

Between the variable for the first and the last appearance of the respective establishment ID ('grd_dat' and 'lzt_dat') in the core dataset and the information on the foundation and closure of

establishments in the extension file 'entry and exit' there are conceptual differences and differences concerning data source.

In order to create the extension file 'entry and exit' a social security notification is first required and then additional information on worker flows into the newly founded establishment or from the closed establishment, respectively. The following limitations result:

- Classification of establishment start-ups in western Germany since 1976, in eastern Germany since 1993

- Classification of establishment closures one year before the last available data

Though the 'worker flows' dataset has the same data source (BeH) and the same reference dates, only one notification is considered for each social security number. For this, the main occupation is used. Possible additional occupations are not used. In order to obtain a time-consistent definition, marginal part-time jobs are not considered. This has the following consequences:

- For establishments that initially report only employees in marginal part-time employment, the year of foundation shifts to the year in which the first non-marginal job was reported.

- For establishments that report only employees in marginal part-time employment in the final years, the year of closure shifts to the year in which a non-marginal job was last reported.

- For establishments that initially consist solely of employees whose job in that establishment is not their main occupation, the year of foundation shifts to the year in which the first main occupation was reported.

- For establishments that consist in the final years solely of employees whose job in that establishment is not their main occupation, the year of closure shifts to the year in which a main occupation was last reported.

- The establishment size in the extension file is equivalent to the number of employees whose main occupation is in the respective establishment, without marginal part-time employees in the year of foundation or the year of closure; this figure may differ from the total number of employees (az_ges) and from the difference between the total number of employees and the number of people in marginal part-time employment.

The information on entries and exits is an extension file of the BHP which is only made available on application.

## 3.6    Generating the extension file – bankruptcies

Many empirical studies about firm exits point out that it is important to distinguish between different types of closures, e.g., voluntary and involuntary liquidations. This extension file offers a way to identify exits due to bankruptcies in the BHP from 2007 onwards. In contrast to other closures, bankruptcies can be unambiguously regarded as indication for economic failure and can therefore be interpreted as involuntary exits.

It should however be noted that the data do not allow to construct a complete bankruptcy statistic (which is available at the German Federal Statistical Office) for mainly two reasons. First, the BHP contains only establishments with employees subject to social security. Bankruptcies of firms without employees are therefore not visible. Second, the information in the BHP refers to establishments, i.e., local units, rather than firms (legal units), whereas bankruptcy occurs at the

firm level. Hence, if a multi-plant firm goes bankrupt our data still captures all establishments belonging to this firm, but we cannot count the number of bankrupt firms and we do not know which establishments belong to the same firm.

A detailed description of an earlier version of the data set can be found in Fackler et al. (2017), which can be downloaded free-of-charge from the FDZ homepage. All users who wish to work with the extension file 'bankruptcies' are recommended to read this Methodenreport.

The information on bankruptcies is an extension file of the BHP, which is only made available on application.

### 3.6.1    Data sources

The bankruptcy data set is obtained from three data sources. The largest part is collected routinely by the BA's local branches in the course of the administration of Insolvenzgeld, a compensation scheme for workers who do not receive their wages due to their employer's bankruptcy. In general, every employee is eligible, even employees who are not subject to social security. Insolvenzgeld is provided by the BA and each case is administered by one of the approximately 600 local branches of the BA. Two of our data sources originate from the administration of Insolvenzgeld. This information is complemented with data from another source, namely social security notifications. In what follows we describe each source in more detail.

First, the directives of the BA on the implementation of Insolvenzgeld require close local monitoring. Therefore, the employees of the local branches of the BA are required to not only handle applications for Insolvenzgeld but also to actively monitor local bankruptcy processes. In particular, they are obligated to be up to date regarding the local situation and to communicate regularly with the courts that decide on the bankruptcy filings. If (upcoming) bankruptcies become known to the BA, the staff that is responsible for Insolvenzgeld stores information on the status of each case (e.g., formal opening of the bankruptcy process, rejection due to lack of assets, or the final closure of the firm; see §165 Social Code III). Importantly, this information is collected even if no employee applies for Insolvenzgeld. Overall, the implementation directives require highly proactive behaviour of the respective BA staff. Moreover, the administration is legally obligated to check the prerequisites for an Insolvenzgeld application, e.g., by requesting court decisions. We refer to these data as "source: monitoring" data.

While the first data source is built mainly for internal usage and to support administrative processes within the BA, the second data source stems from the administration of Insolvenzgeld itself. This information is stored at the level of individual Insolvenzgeld applications and may involve one or more employees per establishment. In order to identify establishments that are affected by bankruptcies from these applications, we marked the establishment identifier in each worker's employment biography (using the BeH) that is closest to the date of the respective Insolvenzgeld application. We refer to these data as "source: application for Insolvenzgeld" data.

Third, we consider data from administrative social security notifications. In Germany, every employer is obliged to notify each employee subject to social security at least once a year to the social insurance. Moreover, there are further mandatory notifications during the year, for example if an employment relationship begins or ends. According to current legislation (§8a DEÜV), employers must inform the social insurance if employees are exempted due to employer bankruptcy. Therefore, this data source is more restrictive than the before mentioned sources as

a firm has to actually exempt at least one of its employees from work. We refer to these data as "source: employment notification" data.

### 3.6.2 Merge with BHP

Since we want to distinguish between establishments that exit the market due to bankruptcies and other closures, we combine the information from the data sources with a list of establishment exits. In the following, we consider all establishments as exits if they vanish from the BHP, i.e., they are counted as exits in year t if they appear in the data that year but not in any of the following years. Since the current version of the BHP ends in 2021, it is possible to identify exits up to the year 2020. Combining this list of exiting establishments from the BHP with the bankruptcy data yields 335,617establishments in the population that exited the market due to bankruptcies between 2007 and 2021. See Fackler et al. (2017) for a comparison with the official statistics on bankruptcies of the German Federal Statistical Office.

To be sure, if an establishment identifier vanishes from the data this does not necessarily reflect a "true" closure because establishment identifiers may also alter due to changes of ownership, of legal form or other administrative reasons (see also section 3.5). For example, the disappearance of the establishment ID may also be triggered by the bankruptcy administrator requesting a new establishment ID when beginning his work which might be close to the date reported in our data sources. But a firm may also survive bankruptcy and then close later due to other reasons. Therefore, it is tempting to concentrate on ID's vanishing within a particular time-span around the date available in the bankruptcy data sources. However, a bankruptcy is a legal procedure that may take up to several years and clearly, an establishment may disappear from the BHP at various points in time during this process. Given the length of the process and the various points in time where an establishment might close, the difference between the date of the bankruptcy information and the date of the establishment exit may cover a long time span and it is hardly possible to decide which time span should reasonably be chosen to ensure that the closure was due to bankruptcy.

We therefore provide the year of an establishment's first and last[5] appearance in our bankruptcy data for each of the three sources. This enables the user to vary this time span. To give an impression, Fackler et al. (2017) found that the majority of all observations have a first appearance in the source data that is within one year before or after the last day an establishment appears in the BA data. However, we advise users to check their results for robustness in this regard.

## 3.7 Sampling procedure

Following the validation procedures described in the previous sections and after the data have been aggregated to establishment level, a 50% simple random sample is drawn from all the establishment IDs in the period 1975–2021.

---

[5] The variable monitoring_yr_2 is no longer available since version BHP 7519, because, as a result of a change in the BA's IT systems, mass observations received a new maximum of the date for the records of the data source "monitoring" in 2020, which has no content-related meaning. On a similar problem, see Fackler et al. (2017, p. 5).

## 3.8    Labelling missing values

In the BHP missing values are coded as follows:

**Table 4: Missing values**

| Term | Value | Description |
|------|-------|-------------|
| .z No entry | .z | Values of a variable which are not systematically missing<br>**Example:** missing information on the establishment's district code |
| .n N/A | .n | A variable is not available for a certain period or a certain group of establishments.<br>**Examples**: Information on the establishment's economic activity using the WZ03 code is only available for the period 2003–2008. In all other years the establishments show the value .n for this variable.<br>Establishments with no full-time employees with unknown qualifications show the value .n for the gross daily wage variable for this group of employees. |

Especially in 2011 there are considerable shares of missing values in the data on individuals on which the BHP is based. Due to the aggregation of the data to establishment level, the labelling of these missing values is lost. E.g. for the variable "number of engineers and natural scientists" only individuals with an appropriate occupational category are counted. Individuals without the appropriate occupational category or with no information are not taken into consideration (see Section 4.6).

## 3.9    Use of German and English labels

The BHP data contain both German and English labels. Using the Stata command "`label language en`" or "`label language de`" it is possible to switch between English and German labels.

## 3.10    Calculation of missing variables

In order to minimise the memory requirement of the BHP, some variables that the users can calculate themselves from the available information are not included in the data. The following Stata codes serve as tools for this:

```
*** Number in workforce stock
* Number of men:
generate az_m = az_ges – az_f
* Number of men working full-time:
generate az_m_vz = az_vz - az_f_vz
* Number of men working part-time:
generate az_m_tz = az_tz - az_f_tz
* Number of employees with unknown qualifications:
generate az_uq = az_ges – (az_gq + az_mq + az_hq)
* Number of employees with unknown qualifications, full-time:
generate az_uq_vz az_vz - (az_gq_vz + az_mq_vz + az_hq_vz)
* Number of non-EU foreigners:
generate az_n_eu = az_ges – (az_d_+ az_eu)
* Number of foreigners working full-time:
generate az_a_vz = az_vz - az_d_vz
* Number of other workers:
generate az_sonst = az_ges –(az_reg + az_gf + az_azubi + az_atz)
```

```
*** Means of imputed gross average daily wages
* Mean imputed gross average daily wage for men (working full-time):
generate te_imp_mw_m = ((az_vz*te_imp_mw)-(az_f_vz*te_imp_mw_f)) / (az_vz-az_f_vz) if
az_vz != az_f_vz & az_f_vz > 0
* Exception: establishments with entire workforce composed of men working full-time:
replace te_imp_mw_m = te_imp_mw if az_f_vz == 0
* Recode missings
recode te_imp_mw_m (.=.n)
* Mean imputed gross average daily wage for employees with unknown qualifications, full-
time:
generate help1 = az_gq_vz*te_imp_mw_gq if te_imp_mw_gq != .n
replace help1 = 0 if te_imp_mw_gq == .n
generate help2 = az_mq_vz*te_imp_mw_mq if te_imp_mw_mq != .n
replace help2 = 0 if te_imp_mw_mq == .n
generate help3 = az_hq_vz*te_imp_mw_hq if te_imp_mw_hq != .n
replace help3 = 0 if te_imp_mw_hq == .n
generate te_imp_mw_uq = ((az_vz*te_imp_mw)-(help1 + help2 + help3)) / (az_vz-(az_gq_vz +
az_mq_vz + az_hq_vz))
* Recode missings
recode te_imp_mw_uq (.=.n)
* Mean imputed gross average daily wage for foreigners working full-time:
generate te_imp_mw_a = ((az_vz*te_imp_mw)-(az_d_vz*te_imp_mw_d)) / (az_vz-az_d_vz) if
az_vz != az_d_vz & az_d_vz > 0
* Exception: establishments with entire workforce composed of foreigners working full-
time:
replace te_imp_mw_a = te_imp_mw if az_d_vz == 0
* Recode missings
recode te_imp_mw_a (.=.n)

*** Inflows and outflows
* Inflows, men:
generate ein_m = ein_ges – ein_f
* Inflows, full-time, male:
generate ein_m_vz = ein_vz - ein_f_vz
* Inflows, part-time, male:
generate ein_m_tz = ein_tz – ein_f_tz
* Outflows, male:
generate aus_m = aus_ges – aus_f
* Outflows, full-time, male:
generate aus_m_vz = aus_vz - aus_f_vz
* Outflows, part-time, male:
generate aus_m_tz = aus_tz – aus_f_tz
```

The sum of *az_vz* and *az_tz* is **not** equivalent to the total number of employees *az_ges*, but to the sum of *az_reg* and *az_sonst*, whereby *az_sonst* has to be calculated. The figure for *az_sonst* is the result of *az_ges* – (*az_reg* + *az_gf* + *az_azubi* + *az_atz*).

Due to the rounding of the wage information (see descriptions of the variables in Section 5), implausible values may occur in few cases when calculating the means of the imputed gross average daily wage.

# 4  Data quality

As the BHP is generated by aggregating employment notifications from the BeH, the data quality of the BHP is linked with that of the BeH. The BeH has the following notes.

## 4.1 Eastern Germany

Due to the introduction of the employment notification procedure in the federal states of eastern Germany, the BeH data for eastern Germany can only be assumed to be sufficiently complete from 1993 onwards. Analyses of eastern German establishments should therefore not begin before 1993 even if data for Eastern Germany are already included from 1992 onwards.

## 4.2 Marginal part-time workers

The considerable increase in the number of BeH observations and thus also in the number of establishments from 1999 onwards is due to the introduction of the obligation to submit employment notifications for people in marginal part-time employment from 1 April 1999 onwards.

## 4.3 Under-recording of notifications in the latest available data

Within the employment notification procedure, a certain time lag is unavoidable. Although changes in employment relationships have to be reported immediately and existing employment relationships have to be confirmed annually by 15 April (or by 15 February since the end of 2013) of the following year, some notifications actually arrive years later. The History File of the IAB is not updated continuously, however, but at certain intervals. In general, this is done using files of employment notifications for one particular year, which were submitted 36, 18 or 6 months after the end of the reporting year (e.g. the 18-month file for 2013 can be created in July 2015 at the earliest). Notifications submitted more than three years late are not taken into account at the IAB, which means that a 36-month file shows a 100 % degree of completeness by definition. For generating the set of yearly BHP data for 1975-2018 it was possible to use 36-month files, for 2019 and 2020 18-month files and for 2021 (only) a 6-month file. It can be assumed that the number of establishments is slightly under-recorded for the years 2019 and 2020. A comparison of the 18- and 36-month files for the year 2018 shows that that in the last 18 months of the waiting period an additional 0.2% of establishments and 0.1% of employees were added. It can also be assumed that there are larger gaps for 2021, which makes it advisable to compare the 6- and 18-month files for 2020: for instance, the 18-month file contains 0.4% more employees than the 6-month file. At establishment level the notifications that were submitted late had a slightly stronger effect: after 18 months an additional 1.8% of the establishments are recorded.

## 4.4 Data on earnings

In 1984, a change was made in the employment notification procedure. From that time onwards, one-off payments of gross earned income were reported as part of the annual earnings subject to social security contributions, which leads to an increase in the average daily wage. In particular the proportion of wages and salaries above the upper earnings limit increases considerably from that year onwards (cf. Bender et al. 1996).

## 4.5    Part-time employees

Especially in 1999, a significant increase in notifications of part-time employment can be observed. This is caused both by the actually observed increase in part-time work as well as by the fact that since 1999 employment notifications have generally been completed more correctly.

## 4.6    Change of the occupation code in 2011

Information on vocational training, the occupation/activity performed and the occupational status is transmitted by means of notifications made by the employer in accordance with the Data Collection and Transmission Regulation (DEÜV) using a so-called occupation code. Notifications for periods with an end date later than 30.11.2011 use the new Occupation Code 2010 (for further details see Bertat et al. 2013). The decision to switch to the new occupation code was made by the central organisations of the social security agencies as a number of facts could no longer be recorded in a way that was up-to-date and realistic using the Occupation Code 2003.

The following variables are affected by this change: working time, occupation, occupational status and educational and vocational qualification level. In addition, after the switch to the new occupation code, details are also available about whether an employment relationship is fixed-term and whether a person is employed by a temporary work agency to be hired out to other firms. The most important consequence is the switchover to a new occupational classification. Instead of the previous Classification of Occupations 1988 (Klassifikation der Berufe 1988 (KldB 1988), see Bundesanstalt für Arbeit 1988), the more highly differentiated Classification of Occupations 2010 (KldB 2010) (see Bundesagentur für Arbeit 2021a, 2021) is reported with the new occupation code.

In addition to the changes in the categories for education and occupation, it is to be noted that the new occupation code no longer records any information about the occupational status. Table 5 compares the contents of the new and the old occupation codes.

Moreover, there was a transitional period from December 2011 until May 2012 during which the verification programs of the notification procedure permitted the omission of information on the occupation codes. About 20% of the notifications for this period therefore have no valid information in the occupation code.

For the BHP, which contains reference-date-specific information for the reference date of 30.06, this switchover to the new occupation code means that the 2011 wave is based on notifications using both the old and the new occupation codes. Moreover, for many employees no information was available from the occupation code for 2011.

When compiling the BHP, one aim was to smooth out as far as possible the inconsistencies in the time series that were caused by the switch to the new occupation code and to replace missing values in the year 2011 using imputation procedures.

**Table 5: Contents of the new and the old occupation codes**

| Contents | Values of old occupation code | Values of new occupation code |
|---|---|---|
| Occupation performed | Classification of Occupations 1988 (KldB 1988) | Classification of Occupations 2010 (KldB 2010) |
| Occupational status | Trainees/apprentices<br>Unskilled employees<br>Skilled workers<br>Master craftsmen, foremen<br>Home workers and freelance home workers | Not recorded |
| Part-time | Part-time employees working less than 18 hours<br>Part-time employees working 18 hours or more | Full-time<br>Part-time |
| Education and vocational training | Lower/intermediate secondary school leaving certificate without completed vocational training<br>Lower/intermediate secondary school leaving certificate with completed vocational training<br>Upper secondary school leaving certificate without vocational training<br>Upper secondary school leaving certificate with vocational training<br>Degree from a university of applied sciences (Fachhochschule)<br>University degree<br>Education and training unknown, no statement possible | Not recorded |
| School education (highest school qualification) | Not recorded | None<br>Lower secondary<br>Intermediate school leaving certificate or equivalent school education<br>Upper secondary school leaving certificate (general or subject-specific aptitude for higher education) |
| Vocational education and training (highest vocational qualification) | Not recorded | No vocational qualification<br>Qualification from a recognised vocational training programme<br>Master craftsman / technician – or equivalent qualification from a technical college<br>Bachelor degree<br>Diploma/Magister/Master's degree/state examination<br>Doctorate |
| Employment relationship via employment agency | Not recorded | Yes<br>No |

| Contents | Values of old occupation code | Values of new occupation code |
|---|---|---|
| Fixed-term contract | Not recorded | Permanent<br>Fixed-term |

## 4.6.1 Effects of the changeover

Both the old and the new occupation codes contain not only information about education and vocational training and on the occupation performed but also details on full-time and part-time employment. Whereas the values of the variables on education, vocational training and occupation have changed in the new code, working hours are still recorded according to the old definition. This means that a person is reported as working part-time when their working hours are less than the regular working hours laid down in the collective agreement or by the company. Only the distinction between "mini" and "midi" part-time (i.e., less or more than 18 hours) is dropped in the new occupation code.

After the introduction of the new occupation code, however, a considerable increase in part-time employment is observed. This can be put down to an updating effect: when the new occupation code was introduced employers obviously checked and, if necessary, updated the information they had reported so far (see Bertat et al. 2013). This suggests that part-time employment was under-recorded in the years prior to 2012.

The introduction of the new occupation code can also be seen to have had an effect on the BHP: whereas only 15 % of employees were reported as working part-time in 2010, the percentage rose to 19 % in 2012. The part-time rate among women rose from 26 % in 2010 to 32 % in 2012. These changes also affect the average daily wages for full-time employees reported in the BHP: whereas the average daily wage stood at 69 Euros in 2010, it was already 76 Euros in 2012. It is also noticeable that the share of establishments with no full-time employees increased from 40 % in 2010 to 45 % in 2012.

## 4.6.2 Missing values in 2011

As mentioned above, from December 2011 until May 2012 there was a transitional period during which the occupation code could be left unfilled, which resulted in a substantial increase in the number of missing values. In the 2011 wave of the BHP the details in the following variables are therefore affected by this:

- Employees according to the Blossfeld classification of occupations: *az_bf_agr, az_bf_emb, az_bf_edi, az_bf_evb, az_bf_qmb, az_bf_qdi, az_bf_qvb, az_bf_tec, az_bf_semi, az_bf_ing, az_bf_prof* and *az_bf_man*
- Engineers and natural scientists: *az_ingnat*
- Employees according to level of requirement: *az_niv1, az_niv2, az_niv3* and *az_niv4*
- Agency workers: *az_leih*
- Temporary employees: *az_bfr*

It is possible to distinguish here between variables that could be generated only from information in the new occupation code (*az_niv1, az_niv2, az_niv3, az_niv4, az_leih, az_bfr*) and variables that were formed from both occupation codes by means of recoding (Blossfeld variables, *az_ingnat*).

In the case of the variables in the latter category, the missing values are restricted to the changeover phase. The new variables on level of task requirement, agency work and fixed-term contracts, on the other hand, have no valid values for the entire period before 2011.

The missing values are examined more closely below. The basis for this is all notifications in the BeH that include the reference date of 30.06.2011. We distinguish here between notifications that have no valid values for all of the variables listed above, and notifications that only have missing values for the new variables.

It can be seen that the missing values are not distributed randomly, but occur very differently from region to region. While there are hardly any unfilled occupation codes for Lower Saxony, missing values are more frequent in districts in southern Germany (see Figure 1 and Figure 2). Differences can also be seen in the individual sectors of the economy: whereas in real estate activities only 8 % of the notifications have missing information in all variables, in the water supply sector the figure is 26 % (see Table 6). The shares are correspondingly larger when the new variables are examined. Fluctuations in the values are also found if the missing values are analysed according to establishment size (see Table 7), but they are significantly smaller than the differences between the districts and sectors.

As the BHP only shows the total number of employees in individual groups, the missing values cannot be identified directly in the data. With regard to the variables of vocational education and vocational training and level of task requirement, it is recommendable to compare the total number of employees in the Blossfeld variables or the level of task requirement variables with the total number of employees (*az_ges*). If discrepancies emerge here, it can be assumed that details on the occupation code were missing for some of the employees in the establishment[6]. A comparison of this kind is not possible for engineers and natural scientists, agency workers or employees on fixed-term contracts. It is therefore not recommended to use the variables *az_ingnat, az_leih* and *az_bfr* for analysing the 2011 wave.

In principle, the variables on full-time and part-time employment and on education and vocational training are also affected by missing values. However, it was possible to eliminate these missing values to a large degree by means of imputation procedures (see Sections 3.1.6 and 3.1.4).

---

[6] Whereas in the 2010 wave a difference between the total number of employees in the Blossfeld variables and the total number of employees could only be detected in 5 % of all the establishments, the figure for 2011 is 42 %.

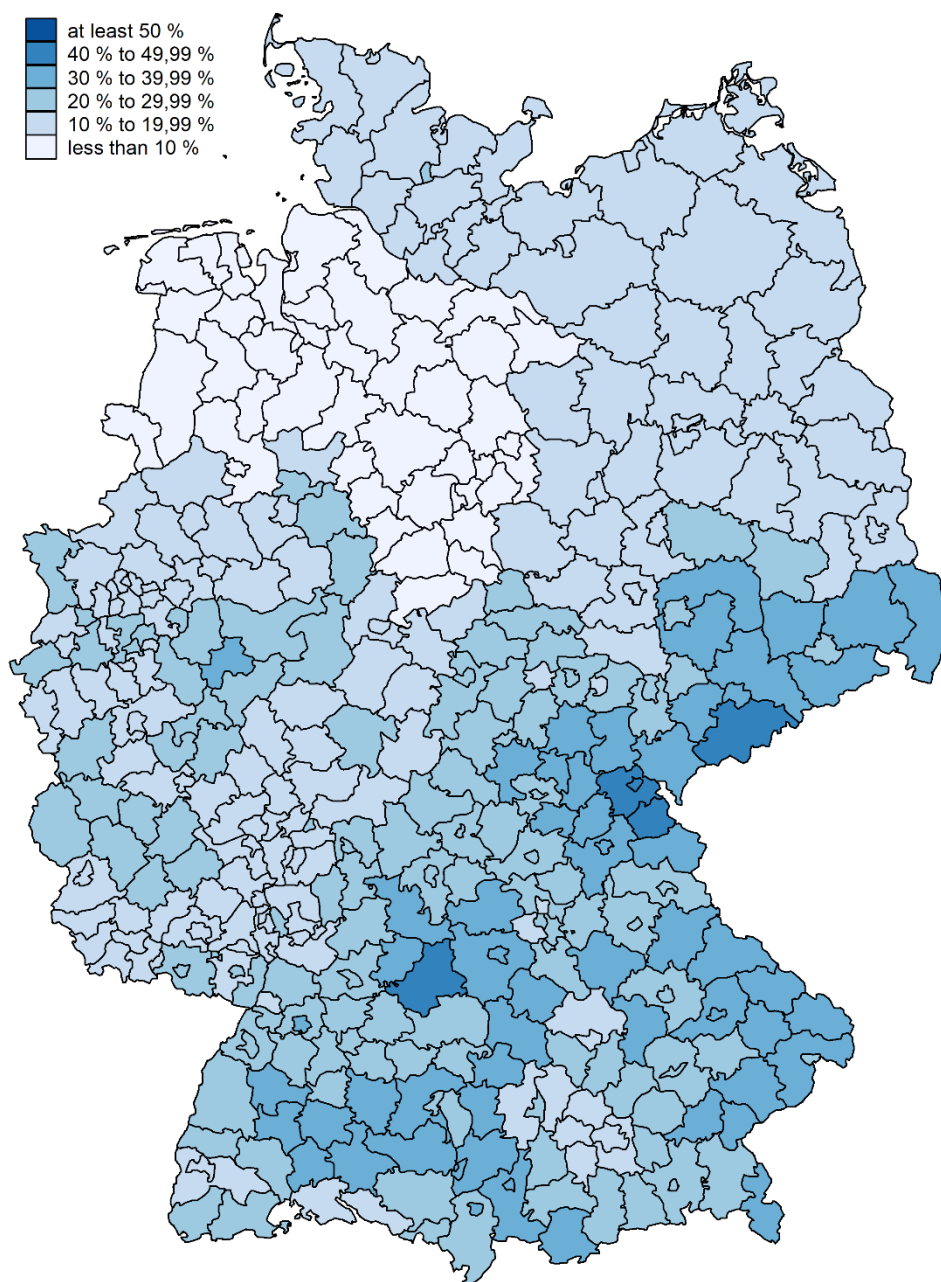**Figure 1: Share of missing values in the occupation code in 2011 by district (all affected variables)**



Legend:
- at least 50 %
- 40 % to 49,99 %
- 30 % to 39,99 %
- 20 % to 29,99 %
- 10 % to 19,99 %
- less than 10 %

**Figure 2: Share of missing values in the occupation code in 2011 by district (only new variables)**



less than 50 %
40 % to 49,99 %
30 % to 39,99 %
20 % to 29,99 %
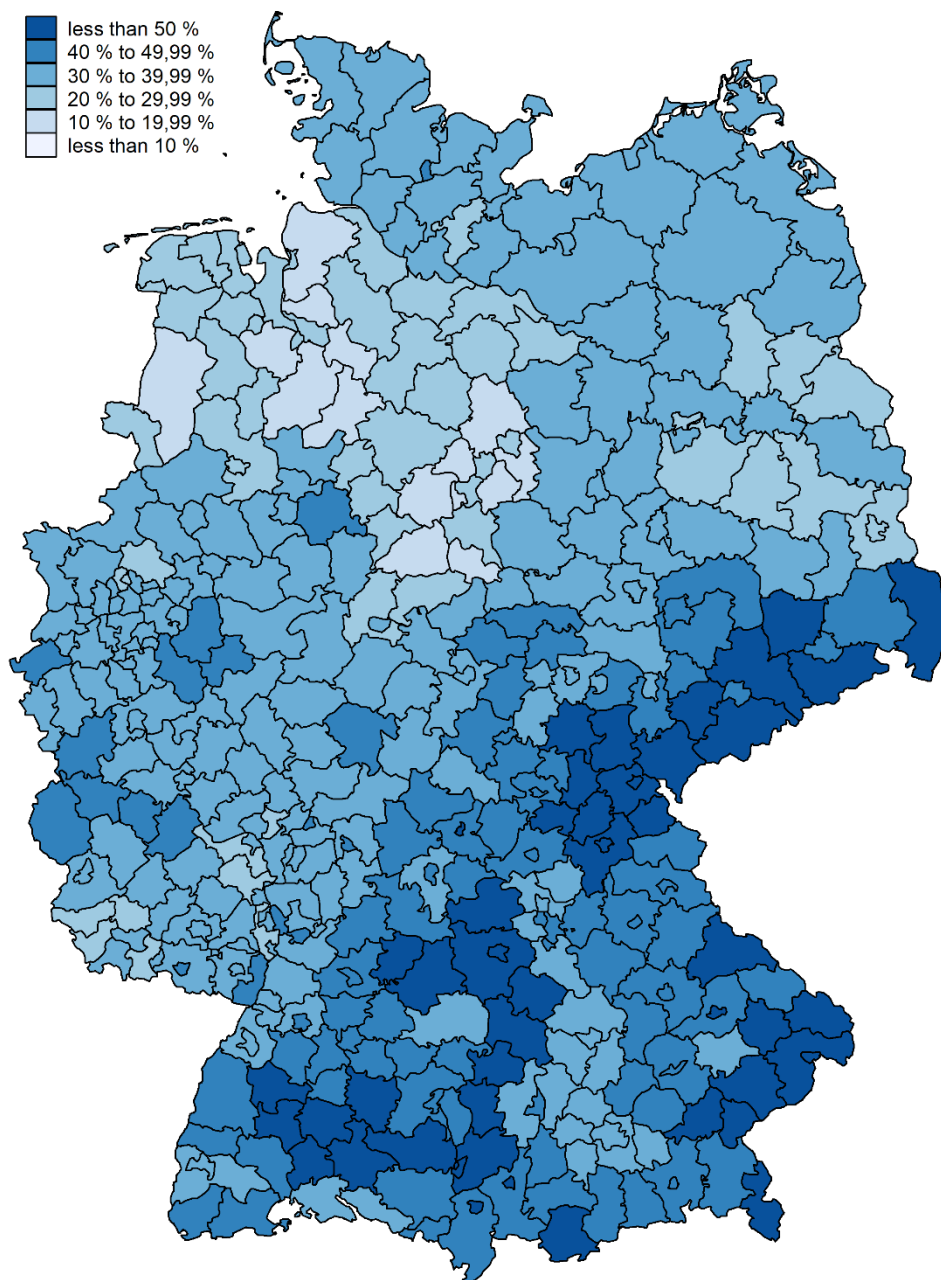10 % to 19,99 %
less than 10 %

**Table 6: Share of missing values in the occupation code in 2011 by sector**

| Sector | Share of missing values in all variables | Share of missing values in the new variables |
|---|---|---|
| Agriculture, hunting and forestry; fishing | 21.02 % | 48.35 % |
| Mining and quarrying | 14.22 % | 26.88 % |
| Manufacturing | 24.03 % | 36.47 % |
| Electricity and gas supply | 15.96 % | 25.98 % |
| Water supply etc. | 26.26 % | 39.09 % |
| Construction | 21.39 % | 40.93 % |
| Wholesale and retail trade; repair of motor vehicles | 16.89 % | 35.06 % |
| Transport and storage | 19.44 % | 37.84 % |
| Hotels and restaurants | 15.15 % | 49.44 % |
| Information and communication | 11.70 % | 28.34 % |
| Financial and insurance services | 12.56 % | 22.50 % |
| Real estate | 8.08 % | 22.17 % |
| Freelance, scientific and technical services | 11.10 % | 27.70 % |
| Other economic services | 17.71 % | 47.04 % |
| Public administration, defence, social security | 19.82 % | 31.14 % |
| Education | 13.43 % | 36.28 % |
| Health and social work | 15.77 % | 31.87 % |
| Art, entertainment, recreational activities | 11.12 % | 39.58 % |
| Other service activities | 13.31 % | 31.42 % |
| Private households | 77.38 % | 90.55 % |
| Extra-territorial organisations and bodies | 32.98 % | 43.02 % |
| Total | 18.88 % | 36.82 % |

Source: BeH V10.00.00; reference date: 30.06.2011; own calculations

**Table 7: Share of missing values in the occupation code in 2011 by establishment size**

| Number of employees in the establishment | Share of missing values in all variables | Share of missing values in the new variables |
|---|---|---|
| 1-4 | 18.04 % | 38.17 % |
| 5-9 | 14.51 % | 35.16 % |
| 10-19 | 16.37 % | 36.74 % |
| 20-49 | 19.22 % | 38.66 % |
| 50-99 | 20.75 % | 39.27 % |
| 100-199 | 21.62 % | 39.11 % |
| 200-499 | 20.86 % | 36.49 % |
| 500 or more | 16.76 % | 31.00 % |
| Total | 18.88 % | 36.82 % |

Source: BeH V10.00.00; reference date: 30.06.2011; own calculations

# 5 Description of the variables and characteristics

Frequency tables and overviews of the individual values and labels of the variables can be found in separate files on the FDZ website (see 8.3 in the Appendix).

## 5.1 Core dataset

### 5.1.1 Establishment characteristics

#### 5.1.1.1 Establishment ID (bhp_betnr)

| Category | Description |
|---|---|
| Variable label | Establishment id |
| Variable name | bhp_betnr |
| Data type | numerical |
| Origin | generated variable |
| Detailed description | To make the data anonymous, the original establishment numbers allocated by the Federal Employment Agency are replaced by randomly generated, yet unambiguous establishment IDs. Further information on the allocation of establishment numbers and on the definition of an establishment can be found in Section 2.2.<br>The artificial establishment ID may be used to merge the yearly data of the core dataset to create a panel dataset and to add the extension files and the file with the time-consistent classifications of economic activities to the sets of yearly data of the core dataset. |
| Sensitive variable | no |

#### 5.1.1.2 Year of first appearance (grd_dat)

| Category | Description |
|---|---|
| Variable label | Date of first appearance |
| Variable name | grd_dat |
| Data type | date |
| Origin | BeH |
| Detailed description | The variable states the precise date on which the respective establishment ID first appears in the BeH data. If an establishment number is first indicated after 1975 for western Germany, or after 1990 for eastern Germany, it can be presumed that the variable represents the year in which the establishment in question was founded. However, it could also be an older establishment that has been allocated a new establishment number after a change of ownership or legal form. The employment agencies also allocate new establishment numbers in some cases when parts of companies are outsourced (for more detailed information see Bundesagentur für Arbeit (2007: 9–11)). It is also possible that the establishment existed previously but had no employees covered by social security or from 1999 onwards also no marginal part-time employees.<br>An establishment is not necessarily contained in the BHP in the year of its |

| Category | Description |
|---|---|
| | first appearance, as only the reference date of 30.06 is relevant for the BHP. If the establishment has no employees on this reference date in the year it was founded, it does not appear in the BHP.<br><br>The first appearance of an establishment can be identified more precisely using the variable 'eintritt' in the Extension File 'entry and exit'.<br><br>Further information on the allocation of establishment numbers and on the definition of an establishment can be found in Section 2.2. |
| Sensitive variable | no |

### 5.1.1.3 Year of last appearance (lzt_dat)

| Category | Description |
|---|---|
| Variable label | Date of last appearance |
| Variable name | lzt_dat |
| Data type | date |
| Origin | BeH |
| Detailed description | The variable states the precise date on which the respective establishment ID appears in the BeH data for the last time. (cf. Bender et al. 1996).<br><br>If the existence of an establishment number in the BHP ends before 2019, this could be a case of closure of the establishment. However, other possible causes are also a "random change of establishment number on change of ownership or legal form", an "outsourcing of parts of the company under a new number" or other administrative changes. For more detailed information, see Bender et al. 1996: 15f. and pp. 27–30 or Bundesagentur für Arbeit 2007: 9–11).<br><br>An establishment is not necessarily contained in the BHP in the year of its last appearance, as only the reference date of 30.06 is relevant for the BHP. If the establishment has no employees on this reference date in the year of its closure, it does not appear in the BHP.<br><br>The last appearance of an establishment can be identified more precisely using the variable 'austritt' in the Extension File 'entry and exit'.<br><br>Further information on the allocation of establishment numbers and on the definition of an establishment can be found in Section 2.2. |
| Sensitive variable | no |

### 5.1.1.4 District code (ao_kreis)

| Category | Description |
|---|---|
| Variable label | Place of work - district (Kreis) |
| Variable name | ao_kreis |
| Data type | Numerical |
| Hierarchy | Federal state \| district |
| Origin | BeH |
| Detailed description | The district code for the workplace originates from the BeH. Here, the district (urban district or rural district) in which the establishment is located is specified. The 5-digit district code contains the following: the first two digits refer to the code for the German federal state (*Bundesland*, NUTS-1), the first to third digits refer to the administrative district (*Regierungsbezirk*, NUTS-2) and the first to fifth digits the district (*Kreis*, NUTS-3). In states with no administrative district, the third digit is 0. |

| Category | Description |
|---|---|
| | To ensure consistent regional allocations across the entire observation period, the district data were recoded to the territorial allocations as of 31.12.2020, i.e., the allocation of an establishment location to a district in all calendar years is based on the boundaries existing as of 31.12.2020. As the district borders changed over time, without territorial allocation updates cases would arise in which the district code of the establishment location changes without the establishment having relocated. |
| Sensitive variable | yes |
| Notes | As a standard, only the German state (ao_bula) is made available as regional information. |

### 5.1.1.5 German state (ao_bula)

| Category | Description |
|---|---|
| Variable label | Place of work - federal state (Bundesland) |
| Variable name | ao_bula |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable indicates the German state in which the establishment is located. It is generated from the district code (ao_kreis). The first two digits of the district code indicate the federal state (*Bundesland*, NUTS-1). |
| Sensitive variable | no |

### 5.1.1.6 Classification of economic activities 73 (w73_3)

| Category | Description |
|---|---|
| Variable label | Classification of economic activities 73 |
| Variable name | w73_3 |
| Data type | Numerical |
| Hierarchy | Economic sector (1-digit code) \| economic group (2-digit code) \| economic class ( 3-digit code) |
| Origin | BeH |
| Detailed description | This variable identifies the economic activity as a 3-digit code in accordance with the WS73 classification. The information is available from 1975 up to and including 2002. WS73 stands for the "Industrial Classification of Economic Activities for the Statistical Office of the Federal Employment Agency, 1973 Edition" (see Bundesanstalt für Arbeit 1973). 269 classes of activity are differentiated by means of a 3-digit code, whereby the first digit defines the economic sector, of a total of 10, and the first two digits together determine the respective group, of a total of 95. Only one code may be allocated to each establishment. Establishments are assigned to the relevant economic class on the basis of their institutional orientation. The assignment of an establishment to an economic class may change over time, however. |
| Sensitive variable | no |

### 5.1.1.7 Classification of economic activities 93, groups (w93_3)

| Category | Description |
|---|---|
| Variable label | Classification of economic activities 93, groups |
| Variable name | w93_3 |
| Data type | Numerical |

| Category | Description |
|---|---|
| Hierarchy | Economic section (1-digit code) \| Economic division (2-digit code) \| Economic group (3-digit code) \| Economic class (4-digit code) \| Economic sub-class (5-digit code) |
| Origin | BeH |
| Detailed description | This variable identifies the economic activity as a 3-digit code in accordance with the WZ93 classification. The information is available from 1999 up to and including 2003. WZ93 stands for the "Industrial Classification of Economic Activities for the Statistical Office of the Federal Employment Agency, 1993 Edition" (see Bundesanstalt für Arbeit 1996). The WZ93 is based on the statistical system of economic activities in the European Community, NACE Rev.1 („Nomenclature génerale des activités économiques dans les communautés européennes"),which has four structural levels, the first two of which are based in turn on the international standard ISIC Rev.3 ("International Standard Industrial Classification of All Economic Activities"). |
| | Only one code may be allocated to each establishment; the classification can be used in its full depth structure, or in a shortened form as necessary. If an establishment operates in various different economic areas, the main economic activity, i.e. the economic focus, has to be established. |
| | The assignment of an establishment to an economic class may change over time, however. |
| Sensitive variable | no |

### 5.1.1.8 Classification of economic activities 93, sub-classes (w93_5)

| Category | Description |
|---|---|
| Variable label | Classification of economic activities 93, sub-classes |
| Variable name | w93_5 |
| Data type | Numerical |
| Hierarchy | Economic section (1-digit code) \| Economic division (2-digit code) \| Economic group (3-digit code) \| Economic class (4-digit code) \| Economic sub-class (5-digit code) |
| Origin | BeH |
| Detailed description | This variable identifies the economic activity as a 5-digit code in accordance with the WS93 classification. The information is available from 1999 up to and including 2003. WZ93 stands for the "Industrial Classification of Economic Activities for the Statistical Office of the Federal Employment Agency, 1993 Edition" (see Bundesanstalt für Arbeit 1996). The WZ93 is based on the statistical system of economic activities in the European Community, NACE Rev.1 („Nomenclature génerale des activités économiques dans les communautés européennes"),which has four structural levels, the first two of which are based in turn on the international standard ISIC Rev.3 ("International Standard Industrial Classification of All Economic Activities"). |
| | Only one code may be allocated to each establishment; the classification can be used in its full depth structure, or in a shortened form as necessary. If an establishment operates in various different economic areas, the main economic activity, i.e. the economic focus, has to be established. |

| Category | Description |
|---|---|
| | The assignment of an establishment to an economic class may change over time, however. |
| Sensitive variable | no |
| Notes | The 3-digit classification of economic activities (w93_3) is made available as a standard. |

### 5.1.1.9 Classification of economic activities 03, groups (w03_3)

| Category | Description |
|---|---|
| Variable label | Classification of economic activities 03, groups |
| Variable name | w03_3 |
| Data type | Numerical |
| Hierarchy | Economic section (1-digit code) | Economic division (2-digit code) | Economic group (3-digit code) | Economic class (4-digit code) | Economic sub-class (5-digit code) |
| Origin | BeH |
| Detailed description | This variable identifies the economic activity as a 3-digit code in accordance with the WZ03 classification. The information is available from 2003 up to and including 2008. WZ03 stands for the "Industrial Classification of Economic Activities, 2003 Edition" of the Federal Statistical Office (see Statistisches Bundesamt 2003). Like the WZ93, the WZ03 is based on the statistical system of economic activities in the European Community, NACE Rev.1 (for more information on NACE Rev. 1 see variable description w93_3, w93_5). The Classifications of Economic Activities have been updated, whereby the structure of WZ93 has largely been retained. |
| | Only one code may be allocated to each establishment; the classification can be used in its full depth structure, or in a shortened form as necessary. If an establishment operates in various different economic areas, the main economic activity, i.e. the economic focus, has to be established. |
| | The assignment of an establishment to an economic class may change over time, however. |
| Sensitive variable | no |

### 5.1.1.10 Classification of economic activities 03, sub-classes (w03_5)

| Category | Description |
|---|---|
| Variable label | Classification of economic activities 03, sub-classes |
| Variable name | w03_5 |
| Data type | Numerical |
| Hierarchy | Economic section (1-digit code) |Economic division (2-digit code) | Economic group (3-digit code) | Economic class (4-digit code) | Economic sub-class (5-digit code) |
| Origin | BeH |
| Detailed description | This variable identifies the economic activity as a 5-digit code in accordance with the WZ03 classification. The information is available from 2003 up to and including 2008. WZ03 stands for the "Industrial Classification of Economic Activities, 2003 Edition" of the Federal Statistical Office (see Statistisches Bundesamt 2003). Like the WZ93, the WZ03 is based the statistical system of economic activities in the |

| Category | Description |
|---|---|
| | European Community, NACE Rev.1 (for more information on NACE Rev. 1 see variable description w93_3, w93_5). The Classifications of Economic Activities have been updated, whereby the structure of WZ93 has largely been retained.<br><br>Only one code may be allocated to each establishment; the classification can be used in its full depth structure, or in a shortened form as necessary. If an establishment operates in various different economic areas, the main economic activity, i.e. the economic focus, has to be established.<br><br>The assignment of an establishment to an economic class may change over time, however. |
| Sensitive variable | Yes |
| Notes | The 3-digit classification of economic activities (w03_3) is made available as a standard. |

### 5.1.1.11 Classification of economic activities 08, groups (w08_3)

| Category | Description |
|---|---|
| Variable label | Classification of economic activities 08, groups |
| Variable name | w08_3 |
| Data type | Numerical |
| Hierarchy | Economic section (1-digit code) \| Economic division (2-digit code) \| Economic group (3-digit code) \| Economic class (4-digit code) \| Economic sub-class (5-digit code) |
| Origin | BeH |
| Detailed description | This variable identifies the economic activity as a 3-digit code in accordance with the WZ08 classification. The information is available from 2008 onwards. WZ08 stands for the "Industrial Classification of Economic Activities, 2008 Edition" of the Federal Statistical Office (see Statistisches Bundesamt 2009). The WZ08 is based on the statistical system of economic activities in the European Community, NACE Rev.2.<br><br>Only one code may be allocated to each establishment; the classification can be used in its full depth structure, or in a shortened form as necessary. If an establishment operates in various different economic areas, the main economic activity, i.e. the economic focus, has to be established.<br><br>The assignment of an establishment to an economic class may change over time, however. |
| Sensitive variable | No |

### 5.1.1.12 Classification of economic activities 08, sub-classes (w08_5)

| Category | Description |
|---|---|
| Variable label | Classification of economic activities 08, sub-classes |
| Variable name | w08_5 |
| Data type | Numerical |
| Hierarchy | Economic section (1-digit code) \| Economic division (2-digit code) \| Economic group (3-digit code) \| Economic class (4-digit code) \| Economic sub-class (5-digit code) |
| Origin | BeH |

| Category | Description |
|---|---|
| Detailed description | This variable identifies the economic activity as a 5-digit code in accordance with the WZ08 classification. The information is available from 2008 onwards. WZ08 stands for the "Industrial Classification of Economic Activities, 2008 Edition" of the Federal Statistical Office (see Statistisches Bundesamt 2009). The WZ08 is based on the statistical system of economic activities in the European Community, NACE Rev.2. Only one code may be allocated to each establishment; the classification can thus be used in its full depth structure, or in a shortened form as necessary. If an establishment operates in various different economic areas, the main economic activity, i.e. the economic focus, has to be established.<br>The assignment of an establishment to an economic class may change over time, however. |
| Sensitive variable | Yes |
| Notes | The 3-digit classification of economic activities (w08_3) is made available as a standard. |

## 5.1.2 General employment structure

### 5.1.2.1 No. employees (az_ges)

| Category | Description |
|---|---|
| Variable label | No. employees total |
| Variable name | az_ges |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees of an establishment for whom a social security notification exists on 30 June of a year. Since the introduction of the new employment notification regulations in 1999, marginal part-time employees have also been included. The variable dormant employment relationships is not included (average daily wage = 0). |
| Sensitive variable | No |

### 5.1.2.2 No. female employees (az_f)

| Category | Description |
|---|---|
| Variable label | No. female employees |
| Variable name | az_f |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of female employees in an establishment. |
| Sensitive variable | No |
| Notes | The gender is determined at individual level by comparing information from different sources (e.g. social security notifications, jobseeker registration at the BA). The aim is to determine a person's gender consistently across all sources. In a few exceptional cases this is incorrect due to errors in data preparation.<br>The number of males is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.2.3 No. regular workers (az_reg)

| Category | Description |
|---|---|
| Variable label | No. regular workers |
| Variable name | az_reg |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of workers in the establishment who are reported with the person group codes 101, 113, 114, 140 and 143. Further information on the person group code can be found in the Appendix (Section 8.2). |
| Sensitive variable | No |
| Notes | The number of "other workers" is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.2.4 No. marginal part-time workers (az_gf)

| Category | Description |
|---|---|
| Variable label | No. marginal part-time workers |
| Variable name | az_gf |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The number of marginal part-time employees is generated from the person group code – values 109 and 209. Further information on the person group code can be found in the Appendix (Section 8.2). This variable only contains valid values in the dataset since 1999 as people in marginal part-time employment were only integrated into the social security notification procedure from that year onwards. |
| Sensitive variable | No |
| Notes | The number of "other workers" is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.2.5 No. trainees/apprentices (az_azubi)

| Category | Description |
|---|---|
| Variable label | No. trainees/apprentices |
| Variable name | az_azubi |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains all employees who are identified as trainees/ apprentices in the person group code (102, 121, 122, 141, 144). Further information on the person group code can be found in the Appendix (Section 8.2). |
| Sensitive variable | No |
| Notes | As this person group has only been reported since 1999, the variable was generated artificially for the time before 1999. Inconsistencies may occur in the time series here.<br>The number of "other workers" is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.2.6 No. employees in partial retirement (az_atz)

| Category | Description |
|---|---|
| Variable label | No. employees in partial retirement |

| Category | Description |
|---|---|
| Variable name | az_atz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of employees in the establishment who are reported with the person group codes 103 and 142. Further information on the person group code can be found in the Appendix (Section 8.2). Partial retirement was introduced in Germany in 1989 and the statutory provisions were amended again as of 01.08.1996. A considerable increase in the number of individuals in partial retirement can therefore be seen in the data from 1997 onwards. |
| Sensitive variable | No |
| Notes | The number of "other workers" is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.2.7 No. full-time (regular workers + others) (az_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time (regular workers + others) |
| Variable name | az_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of individuals in the establishment who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |

### 5.1.2.8 No. part-time (regular workers + others) (az_tz)

| Category | Description |
|---|---|
| Variable label | No. part-time (regular workers + others) |
| Variable name | az_tz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of individuals in the establishment who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as part-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |

### 5.1.2.9 No. full-time female employees (az_f_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time female employees |
| Variable name | az_f_vz |

| Category | Description |
|---|---|
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of females in the establishment who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |
| Notes | The number of full-time male employees is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.2.10 No. part-time female employees (az_f_tz)

| Category | Description |
|---|---|
| Variable label | No. part-time female employees |
| Variable name | az_f_tz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of females in the establishment who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as part-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | no |
| Notes | The number of part-time male employees is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.2.11 No. full-time regular workers (az_reg_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time regular workers |
| Variable name | az_reg_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of individuals in the establishment who are reported with the person group codes 101, 113, 114, 140 and 143 and as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. |
| Sensitive variable | no |

## 5.1.3  Structure of employees by educational and vocational qualifications

### 5.1.3.1  No. unskilled employees (az_gq)

| Category | Description |
|---|---|
| Variable label | No. unskilled employees |

| Category | Description |
|---|---|
| Variable name | az_gq |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable encompasses all employees of an establishment who have a low skill level, i.e. individuals with a lower secondary, intermediate secondary or upper secondary school leaving certificate but no vocational qualifications.<br>In order to replace missing values and to obtain a time-consistent classification, the details on educational and vocational qualifications in the underlying data on individuals were recoded and corrected (see Section 3.1.4). |
| Sensitive variable | no |
| Notes | The number of employees with unknown qualifications is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.3.2  No. qualified employees (az_mq)

| Category | Description |
|---|---|
| Variable label | No. qualified employees |
| Variable name | az_mq |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable encompasses all employees of an establishment who have a medium skill level, i.e. individuals with a lower secondary, intermediate secondary or upper secondary school leaving certificate and a vocational qualification.<br>In order to replace missing values and to obtain a time-consistent classification, the details on educational and vocational qualifications in the underlying data on individuals were recoded and corrected (see Section 3.1.4). |
| Sensitive variable | no |
| Notes | The number of employees with unknown qualifications is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.3.3  No. highly qualified employees (az_hq)

| Category | Description |
|---|---|
| Variable label | No. highly qualified employees |
| Variable name | az_hq |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable encompasses all employees of an establishment who have a degree from a university of applied sciences (Fachhochschule) or a university.<br>In order to replace missing values and to obtain a time-consistent classification, the details on educational and vocational qualifications in the underlying data on individuals were recoded and corrected (see Section 3.1.4). |
| Sensitive variable | No |
| Notes | The number of employees with unknown qualifications is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.3.4 No. full-time and unskilled employees (az_gq_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time and unskilled employees |
| Variable name | az_gq_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable encompasses all full-time employees of an establishment who have a low skill level, i.e. individuals with a lower secondary, intermediate secondary or upper secondary school leaving certificate but no vocational qualifications.<br>In order to replace missing values and to obtain a time-consistent classification, the details on educational and vocational qualifications in the underlying data on individuals were recoded and corrected (see Section 3.1.4).<br>Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |
| Notes | The number of full-time employees with unknown qualifications is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.3.5 No. full-time and qualified employees (az_mq_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time and qualified employees |
| Variable name | az_mq_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable encompasses all full-time employees of an establishment who have a medium skill level, i.e. individuals with a lower secondary, intermediate secondary or upper secondary school leaving certificate and a vocational qualification.<br>In order to replace missing values and to obtain a time-consistent classification, the details on educational and vocational qualifications in the underlying data on individuals were recoded and corrected (see Sections 3.1.4)<br>Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |

| Category | Description |
|---|---|
| Notes | The number of full-time employees with unknown qualifications is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.3.6 No. full-time and highly qualified employees (az_hq_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time and highly qualified employees |
| Variable name | az_hq_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable encompasses all full-time employees of an establishment who have a degree from a university of applied sciences (Fachhochschule) or a university. <br><br> In order to replace missing values and to obtain a time-consistent classification, the details on educational and vocational qualifications in the underlying data on individuals were recoded and corrected (see Section 3.1.4). <br><br> Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |
| Notes | The number of full-time employees with unknown qualifications is not included in the BHP but can be calculated by the user (see Section 3.10). |

## 5.1.4 Employee age structure

### 5.1.4.1 No. employees aged 15–19 (az_15_19)

| Category | Description |
|---|---|
| Variable label | No. employees aged 15-19 |
| Variable name | az_15_19 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in the establishment who are aged between 15 and 19. <br><br> The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. |
| Sensitive variable | No |

### 5.1.4.2 No. employees aged 20–24 (az_20_24)

| Category | Description |
|---|---|
| Variable label | No. employees aged 20-24 |
| Variable name | az_20_24 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in the establishment who are aged between 20 and 24. |

| Category | Description |
|---|---|
| | The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. |
| Sensitive variable | No |

### 5.1.4.3   No. employees aged 25–29 (az_25_29)

| Category | Description |
|---|---|
| Variable label | No. employees aged 25-29 |
| Variable name | az_25_29 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in the establishment who are aged between 25 and 29.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. |
| Sensitive variable | No |

### 5.1.4.4   No. employees aged 30–34 (az_30_34)

| Category | Description |
|---|---|
| Variable label | No. employees aged 30-34 |
| Variable name | az_30_34 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in the establishment who are aged between 30 and 34.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. |
| Sensitive variable | No |

### 5.1.4.5   No. employees aged 35–39 (az_35_39)

| Category | Description |
|---|---|
| Variable label | No. employees aged 35-39 |
| Variable name | az_35_39 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in the establishment who are aged between 35 and 39.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. |
| Sensitive variable | No |

### 5.1.4.6   No. employees aged 40–44 (az_40_44)

| Category | Description |
|---|---|
| Variable label | No. employees aged 40-44 |
| Variable name | az_40_44 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in the establishment who are aged between 40 and 44.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. |

| Category | Description |
| --- | --- |
| Sensitive variable | No |

### 5.1.4.7 No. employees aged 45–49 (az_45_49)

| Category | Description |
| --- | --- |
| Variable label | No. employees aged 45-49 |
| Variable name | az_45_49 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in the establishment who are aged between 45 and 49.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. |
| Sensitive variable | No |

### 5.1.4.8 No. employees aged 50–54 (az_50_54)

| Category | Description |
| --- | --- |
| Variable label | No. employees aged 50-54 |
| Variable name | az_50_54 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in the establishment who are aged between 50 and 54.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. |
| Sensitive variable | No |

### 5.1.4.9 No. employees aged 55–59 (az_55_59)

| Category | Description |
| --- | --- |
| Variable label | No. employees aged 55-59 |
| Variable name | az_55_59 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in the establishment who are aged between 55 and 59.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. |
| Sensitive variable | No |

### 5.1.4.10 No. employees aged 60–64 (az_60_64)

| Category | Description |
| --- | --- |
| Variable label | No. employees aged 60-64 |
| Variable name | az_60_64 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in the establishment who are aged between 60 and 64.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. |
| Sensitive variable | No |

### 5.1.4.11 No. employees aged 65 and older (az_ab65)

| Category | Description |
|---|---|
| Variable label | No. employees aged 65 and older |
| Variable name | az_ab65 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in the establishment who are aged 65 or above.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. |
| Sensitive variable | No |

### 5.1.4.12 No. full-time employees aged 15–19 (az_15_19_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time employees aged 15-19 |
| Variable name | az_15_19_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of full-time employees in the establishment who are aged between 15 and 19.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |

### 5.1.4.13 No. full-time employees aged 20–24 (az_20_24_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time employees aged 20-24 |
| Variable name | az_20_24_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of full-time employees in the establishment who are aged between 20 and 24.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |

| Sensitive variable | No |
| --- | --- |

### 5.1.4.14 No. full-time employees aged 25–29 (az_25_29_vz)

| Category | Description |
| --- | --- |
| Variable label | No. full-time employees aged 25-29 |
| Variable name | az_25_29_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of full-time employees in the establishment who are aged between 25 and 29.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |

### 5.1.4.15 No. full-time employees aged 30–34 (az_30_34_vz)

| Category | Description |
| --- | --- |
| Variable label | No. full-time employees aged 30-34 |
| Variable name | az_30_34_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of full-time employees in the establishment who are aged between 30 and 34.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |

### 5.1.4.16 No. full-time employees aged 35–39 (az_35_39_vz)

| Category | Description |
| --- | --- |
| Variable label | No. full-time employees aged 35-39 |
| Variable name | az_35_39_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of full-time employees in the establishment who are aged between 35 and 39. |

| Category | Description |
|---|---|
| | The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |

### 5.1.4.17 No. full-time employees aged 40–44 (az_40_44_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time employees aged 40-44 |
| Variable name | az_40_44_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of full-time employees in the establishment who are aged between 40 and 44.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |

### 5.1.4.18 No. full-time employees aged 45–49 (az_45_49_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time employees aged 45-49 |
| Variable name | az_45_49_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of full-time employees in the establishment who are aged between 45 and 49.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |

| Category | Description |
|---|---|
| Sensitive variable | No |

### 5.1.4.19 No. full-time employees aged 50–54 (az_50_54_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time employees aged 50-54 |
| Variable name | az_50_54_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of full-time employees in the establishment who are aged between 50 and 54. <br> The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |

### 5.1.4.20 No. full-time employees aged 55–59 (az_55_59_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time employees aged 55-59 |
| Variable name | az_55_59_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of full-time employees in the establishment who are aged between 55 and 59. <br> The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |

### 5.1.4.21 No. full-time employees aged 60–64 (az_60_64_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time employees aged 60-64 |
| Variable name | az_60_64_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of full-time employees in the establishment who are aged between 60 and 64. |

| Category | Description |
|---|---|
| | The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |

### 5.1.4.22 No. full-time employees aged 65 and older (az_ab65_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time employees aged 65 and older |
| Variable name | az_ab65_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of full-time employees in the establishment who are aged 65 or above.<br>The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |

### 5.1.4.23 Mean age of the total employees (alter_mw)

| Category | Description |
|---|---|
| Variable label | Mean age of the total of employees |
| Variable name | alter_mw |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the mean age of all the employees in the establishment. The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year.<br>The values are rounded to the second decimal digits. |
| Sensitive variable | No |

### 5.1.4.24 Mean age of all full-time employees (alter_mw_vz)

| Category | Description |
|---|---|
| Variable label | Mean age of all full-time employees |
| Variable name | alter_mw_vz |
| Data type | Numerical |

| Category | Description |
|---|---|
| Origin | BeH |
| Detailed description | This variable contains the mean age of all the full-time employees in the establishment. The age is calculated from the date of birth. It is recalculated for every cross-section as of the reference date of 30.06 of the respective year. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account.<br>The values are rounded to the second decimal digits. |
| Sensitive variable | No |

## 5.1.5 Research and development activities

### 5.1.5.1 No. engineers and natural scientists (az_ingnat)

| Category | Description |
|---|---|
| Variable label | No. engineers and natural scientists |
| Variable name | az_ingnat |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees with a degree from a university or a university of applied sciences (Fachhochschule) (after correcting the education/training details by means of imputation, see Appendix, Section 3.1.4) who are employed as engineers or scientists in the establishment. This is the total number of employees with the value 5 or 6 in the B2 code (education/training) and the occupational classification KldB 1988 601–612 or 883. This information can be used as a proxy to measure the establishment's research and development activities.<br>Both education/training and the occupational classification are reported by the employer in the employment details supplied in the employment notification. |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). |

## 5.1.6 No. atypical workers

### 5.1.6.1 No. agency workers (az_leih)

| Category | Description |
|---|---|
| Variable label | No. agency workers |
| Variable name | az_leih |
| Data type | Numerical |
| Origin | BeH |

| Category | Description |
|---|---|
| Detailed description | The variable contains the number of individuals who are hired out to the establishment by an employment agency. The reporting establishment in this case is always the business that sends workers to hiring companies. |
| Sensitive variable | No |
| Notes | The underlying information has only been reported since 2011. The variable therefore contains only missing values until 2010 and incomplete details in 2011. |

### 5.1.6.2 No. temporary employees (az_bfr)

| Category | Description |
|---|---|
| Variable label | No. temporary employees |
| Variable name | az_bfr |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of employees with a fixed-term employment contract. |
| Sensitive variable | No |
| Notes | The underlying information has only been reported since 2011. The variable therefore contains only missing values until 2010 and incomplete details in 2011. |

## 5.1.7 Structure of employees by nationality

### 5.1.7.1 No. Germans (az_d)

| Category | Description |
|---|---|
| Variable label | No. Germans |
| Variable name | az_d |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable states the number of employees with German nationality in an establishment. |
| Sensitive variable | No |
| Notes | The number of non-EU foreigners is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.7.2 No. full-time German employees (az_d_vz)

| Category | Description |
|---|---|
| Variable label | No. full-time German employees |
| Variable name | az_d_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable states the number of full-time employees with German nationality in an establishment. Individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees are counted as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and |

| Category | Description |
|---|---|
| | people in marginal part-time employment or in partial retirement are not taken into account. |
| Sensitive variable | No |
| Notes | The number of foreign full-time employees is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.7.3  No. EU-Europeans (without Germany) (az_eu)

| Category | Description |
|---|---|
| Variable label | No. EU-Europeans (without Germany) |
| Variable name | az_eu |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable states the number of employees in an establishment with the nationality of a member state of the European Union on the reference date (without Germany).<br>EU member states over time:<br>• Since wave 1975: Belgium, Germany, France, Italy, Luxembourg, Netherlands, Denmark, Ireland, United Kingdom<br>• Since Wave 1981: additionally Greece<br>• Since Wave 1986: additionally Portugal, Spain<br>• Since Wave 1995: additionally Finland, Austria, Sweden<br>• Since Wave 2004: additionally Estonia, Latvia, Lithuania, Malta, Poland, Slovakia, Slovenia, Czech Republic, Hungary, Cyprus<br>• Since Wave 2007: additionally Bulgaria and Romania<br>• Since Wave 2014: additionally Croatia<br>• Since Wave 2020: without United Kingdom |
| Sensitive variable | No |
| Notes | The number of non-EU foreigners is not included in the BHP but can be calculated by the user (see Section 3.10). |

## 5.1.8   Structure of employees by Blossfeld occupational group

### 5.1.8.1  No. agricultural occupations (az_bf_agr)

| Category | Description |
|---|---|
| Variable label | No. agricultural occupations |
| Variable name | az_bf_agr |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in agricultural occupations according to the Blossfeld classification of occupations.<br>In the context of the employment notification the employer indicates the occupation held by the employee in the form of the 3-digit code from the Classification of Occupations (KldB 1988, see Bundesanstalt für Arbeit, 1988). This information is used to recode the data to the Blossfeld classification of occupations. This classifies the occupations into a total of 12 groups on the basis of the level of task requirement for the job held. Detailed information on the Blossfeld classification of occupations and its recoding to the KldB can be found in Blossfeld 1987. |
| Sensitive variable | No |

| Category | Description |
|---|---|
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Sections 4.6 and 3.1.5). |

### 5.1.8.2  No. unskilled manual occupations (az_bf_emb)

| Category | Description |
|---|---|
| Variable label | No. unskilled manual occupations |
| Variable name | az_bf_emb |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in unskilled manual occupations according to the Blossfeld classification of occupations. Detailed information on the Blossfeld classification of occupations can be found in the variable description az_bf_agr. |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Sections 4.6 and 3.1.5). |

### 5.1.8.3  No. unskilled services (az_bf_edi)

| Category | Description |
|---|---|
| Variable label | No. unskilled services |
| Variable name | az_bf_edi |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in unskilled services occupations according to the Blossfeld classification of occupations. Detailed information on the Blossfeld classification of occupations can be found in the variable description az_bf_agr. |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Sections 4.6 and 3.1.5). |

### 5.1.8.4  No. unskilled commercial and admin. occupations (az_bf_evb)

| Category | Description |
|---|---|
| Variable label | No. unskilled commercial and admin. occupations |
| Variable name | az_bf_evb |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in unskilled commercial and administrative occupations according to the Blossfeld classification of occupations. |

| Category | Description |
|---|---|
| | Detailed information on the Blossfeld classification of occupations can be found in the variable description az_bf_agr. |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Sections 4.6 and 3.1.5). |

### 5.1.8.5 No. skilled manual occupations (az_bf_qmb)

| Category | Description |
|---|---|
| Variable label | No. skilled manual occupations |
| Variable name | az_bf_qmb |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in skilled manual occupations according to the Blossfeld classification of occupations. Detailed information on the Blossfeld classification of occupations can be found in the variable description az_bf_agr. |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Sections 4.6 and 3.1.5). |

### 5.1.8.6 No. skilled services (az_bf_qdi)

| Category | Description |
|---|---|
| Variable label | No. skilled services |
| Variable name | az_bf_qdi |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in skilled services occupations according to the Blossfeld classification of occupations. Detailed information on the Blossfeld classification of occupations can be found in the variable description az_bf_agr. |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Sections 4.6 and 3.1.5). |

### 5.1.8.7 No. skilled commercial and admin. occupations (az_bf_qvb)

| Category | Description |
|---|---|
| Variable label | No. skilled commercial and admin. occupations |
| Variable name | az_bf_qvb |
| Data type | Numerical |
| Origin | BeH |

| Category | Description |
|---|---|
| Detailed description | This variable contains the number of employees in skilled commercial and administrative occupations according to the Blossfeld classification of occupations.<br>Detailed information on the Blossfeld classification of occupations can be found in the variable description az_bf_agr. |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Sections 4.6 and 3.1.5). |

### 5.1.8.8  No. technicians (az_bf_tec)

| Category | Description |
|---|---|
| Variable label | No. technicians |
| Variable name | az_bf_tec |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employed technicians according to the Blossfeld classification of occupations.<br>Detailed information on the Blossfeld classification of occupations can be found in the variable description az_bf_agr. |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Sections 4.6 and 3.1.5). |

### 5.1.8.9  No. semiprofessions (az_bf_semi)

| Category | Description |
|---|---|
| Variable label | No. semiprofessions |
| Variable name | az_bf_semi |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in semiprofessions according to the Blossfeld classification of occupations.<br>Detailed information on the Blossfeld classification of occupations can be found in the variable description az_bf_agr. |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Sections 4.6 and 3.1.5). |

### 5.1.8.10 No. engineers (az_bf_ing)

| Category | Description |
|---|---|
| Variable label | No. engineers |
| Variable name | az_bf_ing |

| Category | Description |
|---|---|
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employed engineers according to the Blossfeld classification of occupations.<br>Detailed information on the Blossfeld classification of occupations can be found in the variable description az_bf_agr. |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Sections 4.6 and 3.1.5). |

### 5.1.8.11 No. professions (az_bf_prof)

| Category | Description |
|---|---|
| Variable label | No. professions |
| Variable name | az_bf_prof |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees in professions according to the Blossfeld classification of occupations.<br>Detailed information on the Blossfeld classification of occupations can be found in the variable description az_bf_agr. |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Sections 4.6 and 3.1.5). |

### 5.1.8.12 No. managers (az_bf_man)

| Category | Description |
|---|---|
| Variable label | No. managers |
| Variable name | az_bf_man |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employed managers according to the Blossfeld classification of occupations.<br>Detailed information on the Blossfeld classification of occupations can be found in the variable description az_bf_agr. |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Sections 4.6 and 3.1.5). |

### 5.1.9 Structure of employees by level of requirement

#### 5.1.9.1 No. employees unskilled/semiskilled tasks (az_niv1)

| Category | Description |
|---|---|
| Variable label | No. employees unskilled/semiskilled tasks |
| Variable name | az_niv1 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of people who perform unskilled and semi-skilled tasks. The introduction of the Classification of Occupations 2010 (KldB 2010) has made it possible to differentiate systematically between occupational activities according to four degrees of complexity. The level of requirement is given in the last digit of the 5-digit occupation code (see Paulus and Matthes 2013a, 2013b). |
| Sensitive variable | No |
| Notes | The underlying information has only been reported since 2011. The variable therefore contains only missing values until 2010 and incomplete details in 2011 (see Section 4.6). |

#### 5.1.9.2 No. employees skilled tasks (az_niv2)

| Category | Description |
|---|---|
| Variable label | No. employees skilled tasks |
| Variable name | az_niv2 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of people who perform skilled tasks. The introduction of the Classification of Occupations 2010 (KldB 2010) has made it possible to differentiate systematically between occupational activities according to four degrees of complexity. The level of requirement is given in the last digit of the 5-digit occupation code (see Paulus and Matthes 2013a, 2013b). |
| Sensitive variable | No |
| Notes | The underlying information has only been reported since 2011. The variable therefore contains only missing values until 2010 and incomplete details in 2011 (see Section 4.6). |

#### 5.1.9.3 No. employees complex tasks (az_niv3)

| Category | Description |
|---|---|
| Variable label | No. employees complex tasks |
| Variable name | az_niv3 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of people who perform complex specialist tasks. The introduction of the Classification of Occupations 2010 (KldB 2010) has made it possible to differentiate systematically between occupational activities according to four degrees of complexity. The level of requirement is given in the last digit of the 5-digit occupation code (see Paulus and Matthes 2013a, 2013b). |
| Sensitive variable | No |

| Category | Description |
|---|---|
| Notes | The underlying information has only been reported since 2011. The variable therefore contains only missing values until 2010 and incomplete details in 2011 (see Section 4.6). |

### 5.1.9.4 No. employees highly complex tasks (az_niv4)

| Category | Description |
|---|---|
| Variable label | No. employees highly complex tasks |
| Variable name | az_niv4 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of people who perform highly complex tasks. The introduction of the Classification of Occupations 2010 (KldB 2010) has made it possible to differentiate systematically between occupational activities according to four degrees of complexity. The level of requirement is given in the last digit of the 5-digit occupation code (see Paulus and Matthes 2013a, 2013b). |
| Sensitive variable | No |
| Notes | The underlying information has only been reported since 2011. The variable therefore contains only missing values until 2010 and incomplete details in 2011 (see Section 4.6). |

## 5.1.10  Wage structure of full-time employees

### 5.1.10.1 No. employees with censored wages (az_zens)

| Category | Description |
|---|---|
| Variable label | No. employees with censored wages |
| Variable name | az_zens |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable indicates the number of individuals whose (arithmetical) average daily wage is defined as censored (see Section 3.1.3.3). The upper earnings limits are determined for each year (separately for East and West Germany) as 97.5% of the mode of daily wages of full-time employees rounded to whole Euro amounts. The respective valid upper earnings limits can be found at https://doku.iab.de/fdz/Bemessungsgrenzen_de_en.xls. |
| Sensitive variable | No |
| Notes | In contrast to the imputed earnings information in the BHP (see variables te_imp*) all employees are taken into account here, not only the full-time employees. |

### 5.1.10.2 Mean imp. wage all full-time employees (te_imp_mw)

| Category | Description |
|---|---|
| Variable label | Mean imp. wage all full-time employees |
| Variable name | te_imp_mw |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the mean imputed gross average daily wage of an establishment's full-time employees. It does not contain the wages of |

| Category | Description |
|---|---|
| | people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes. |
| | The values are reported in euros for all years. |
| | In accordance with the regulations for social security notifications, employers must specify the employee's gross wage subject to social security contributions for a given period of time ("Zeitraumentgelt" = wage over a given period). Until the end of 1998, employers were obliged to specify only the gross earnings subject to social security contributions. This meant that the only wages recorded were those above the marginal part-time income threshold and below the upper earnings limit for social security contributions. Since 1999, however, as part of the new notification procedure, wages that are below the marginal part-time income threshold also have to be reported. Gross wages that are above the upper earnings limit, however, continue to be capped at this level. To generate the gross average daily wage, the wage for a given period is divided by the number of calendar days in the period. To calculate the mean, these censored wages were imputed (see Section 3.1.3.3). |
| | The values are rounded to the second decimal digits |
| Sensitive variable | No |

### 5.1.10.3 Median imp. wage all full-time employees (te_imp_med)

| Category | Description |
|---|---|
| Variable label | Median imp. wage all full-time employees |
| Variable name | te_imp_med |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the median imputed gross average daily wage of an establishment's full-time employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes. |
| | The values are reported in euros for all years. |
| | Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw). |
| | The values are rounded to the second decimal digits. |
| Sensitive variable | No |

### 5.1.10.4 P25 imp. wage all full-time employees (te_imp_p25)

| Category | Description |
|---|---|
| Variable label | P25 imp. wage all full-time employees |
| Variable name | te_imp_p25 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the p25 percentile of the imputed gross average daily wage of an establishment's full-time employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes. |
| | The values are reported in euros for all years. |
| | Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw). |
| | The values are rounded to the second decimal digits. |

| Category | Description |
|---|---|
| Sensitive variable | No |

### 5.1.10.5 P75 imp. wage all full-time employees (te_imp_p75)

| Category | Description |
|---|---|
| Variable label | P75 imp. wage all full-time employees |
| Variable name | te_imp_p75 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the p75 percentile of the imputed gross average daily wage of an establishment's full-time employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes.<br>The values are reported in euros for all years.<br>Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw).<br>The values are rounded to the second decimal digits. |
| Sensitive variable | No |

### 5.1.10.6 Mean imp. wage full-time female employees (te_imp_mw_f)

| Category | Description |
|---|---|
| Variable label | Mean imp. wage full-time female employees |
| Variable name | te_imp_mw_f |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the mean imputed gross average daily wage of an establishment's full-time female employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes.<br>The values are reported in euros for all years.<br>Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw).<br>The values are rounded to the second decimal digits. |
| Sensitive variable | No |
| Notes | The mean imputed gross average daily wage of full-time male employees is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.10.7 Median imp. wage full-time female employees (te_imp_med_f)

| Category | Description |
|---|---|
| Variable label | Median imp. wage full-time female employees |
| Variable name | te_imp_med_f |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the median imputed gross average daily wage of an establishment's full-time female employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes.<br>The values are reported in euros for all years.<br>Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw). |

| Category | Description |
|---|---|
| | The values are rounded to the second decimal digits. |
| Sensitive variable | No |

### 5.1.10.8 Median imp. wage full-time male employees (te_imp_med_m)

| Category | Description |
|---|---|
| Variable label | Median imp. Wage full-time male employees |
| Variable name | te_imp_med_m |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the median imputed gross average daily wage of an establishment's full-time male employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes.<br>The values are reported in euros for all years.<br>Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw).<br>The values are rounded to the second decimal digits. |
| Sensitive variable | No |

### 5.1.10.9 Mean imp. wage full-time and unskilled employees (te_imp_mw_gq)

| Category | Description |
|---|---|
| Variable label | Mean imp. wage full-time and unskilled employees |
| Variable name | te_imp_mw_gq |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the mean imputed gross average daily wage of an establishment's full-time unskilled employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes.<br>The values are reported in euros for all years.<br>Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw).<br>The values are rounded to the second decimal digits. |
| Sensitive variable | No |
| Notes | The mean imputed gross average daily wage of full-time employees with unknown qualifications is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.10.10      Median imp. wage full-time and unskilled employees (te_imp_med_gq)

| Category | Description |
|---|---|
| Variable label | Median imp. wage full-time and unskilled employees |
| Variable name | te_imp_med_gq |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the median imputed gross average daily wage of an establishment's full-time unskilled employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes.<br>The values are reported in euros for all years. |

| Category | Description |
|---|---|
|  | Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw). The values are rounded to the second decimal digits. |
| Sensitive variable | No |

### 5.1.10.11 Mean imp. wage full-time and qualified employees (te_imp_mw_mq)

| Category | Description |
|---|---|
| Variable label | Mean imp. wage full-time and qualified employees |
| Variable name | te_imp_mw_mq |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the mean imputed gross average daily wage of an establishment's full-time qualified employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes. The values are reported in euros for all years. Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw). The values are rounded to the second decimal digits |
| Sensitive variable | No |
| Notes | The mean imputed gross average daily wage of full-time employees with unknown qualifications is not shown included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.10.12 Median imp. wage full-time and qualified employees (te_imp_med_mq)

| Category | Description |
|---|---|
| Variable label | Median imp. wage full-time and qualified employees |
| Variable name | te_imp_med_mq |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the median imputed gross average daily wage of an establishment's full-time qualified employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes. The values are reported in euros for all years. Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw). The values are rounded to the second decimal digits. |
| Sensitive variable | No |

### 5.1.10.13 Mean imp. wage full-time and highly qualified employees (te_imp_mw_hq)

| Category | Description |
|---|---|
| Variable label | Mean imp. wage full-time and highly qualified employees |
| Variable name | te_imp_mw_hq |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the mean imputed gross average daily wage of an establishment's full-time highly qualified employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes. |

| Category | Description |
|---|---|
| | The values are reported in euros for all years. |
| | Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw). |
| | The values are rounded to the second decimal digits. |
| Sensitive variable | No |
| Notes | The mean imputed gross average daily wage of full-time employees with unknown qualifications is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.10.14 Median imp. wage full-time and highly qualified employees (te_imp_med_hq)

| Category | Description |
|---|---|
| Variable label | Median imp. wage full-time and highly qualified employees |
| Variable name | te_imp_med_hq |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the median imputed gross average daily wage of an establishment's full-time highly qualified employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes. |
| | The values are reported in euros for all years. |
| | Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw). |
| | The values are rounded to the second decimal digits. |
| Sensitive variable | No |

### 5.1.10.15 Median imp. wage full-time employees with unknown qualifications (te_imp_med_uq)

| Category | Description |
|---|---|
| Variable label | Median imp. wage full-time employees with unknown qualifications |
| Variable name | te_imp_med_uq |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the median imputed gross average daily wage of an establishment's full-time employees with unknown qualifications. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes. |
| | The values are reported in euros for all years. |
| | Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw). |
| | The values are rounded to the second decimal digits |
| Sensitive variable | No |

### 5.1.10.16 Mean imp. wage full-time German employees (te_imp_mw_d)

| Category | Description |
|---|---|
| Variable label | Mean imp. wage full-time German employees |
| Variable name | te_imp_mw_d |
| Data type | Numerical |
| Origin | BeH |

| Category | Description |
|---|---|
| Detailed description | This variable contains the mean imputed gross average daily wage of an establishment's full-time German employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes.<br>The values are reported in euros for all years.<br>Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw).<br>The values are rounded to the second decimal digits. |
| Sensitive variable | No |
| Notes | The mean imputed gross average daily wage of full-time foreign employees is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.1.10.17 Median imp. wage full-time German employees (te_imp_med_d)

| Category | Description |
|---|---|
| Variable label | Median imp. wage full-time German employees |
| Variable name | te_imp_med_d |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the median imputed gross average daily wage of an establishment's full-time German employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes.<br>The values are reported in euros for all years.<br>Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw).<br>The values are rounded to the second decimal digits. |
| Sensitive variable | No |

### 5.1.10.18 Median imp. wage full-time foreign employees (te_imp_med_a)

| Category | Description |
|---|---|
| Variable label | Median imp. wage full-time foreign employees |
| Variable name | te_imp_med_a |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the median imputed gross average daily wage of an establishment's full-time foreign employees. It does not contain the wages of people in marginal part-time employment, trainees/apprentices and people in partial retirement schemes.<br>The values are reported in euros for all years.<br>Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw).<br>The values are rounded to the second decimal digits |
| Sensitive variable | No |

### 5.1.10.19 Sum imp. wages all employees (te_imp_sum)

| Category | Description |
|---|---|
| Variable label | Sum imp. wages all employees |
| Variable name | te_imp_sum |
| Data type | Numerical |

| Category | Description |
|---|---|
| Origin | BeH |
| Detailed description | This variable contains the sum of imputed gross average daily wage of all establishment's employees.<br>To calculate the sum, the censored wages of full-time-employees were imputed (see Section 3.1.3.3).<br>Further information on the generation of the variable can be found in Mean imp. wage all full-time employees (te_imp_mw).<br>The values are rounded to the second decimal digits |
| Sensitive variable | No |

## 5.2 Extension file – Worker flows

### 5.2.1 Establishment variables

#### 5.2.1.1 Establishment ID (bhp_betnr)

| Category | Description |
|---|---|
| Variable label | Establishment id |
| Variable name | bhp_betnr |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | To make the data anonymous, the original establishment numbers allocated by the Federal Employment Agency are replaced by randomly generated, yet unambiguous establishment IDs. Further information on the allocation of establishment numbers and on the definition of an establishment can be found in Section 2.2.<br>The artificial establishment ID may be used to merge the yearly data of the core dataset to create a panel dataset and to add the extension files and the file with the time-consistent classifications of economic activities to the sets of yearly data of the core dataset. |
| Sensitive variable | No |

#### 5.2.1.2 Year (jahr)

| Category | Description |
|---|---|
| Variable label | Year |
| Variable name | jahr |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable shows the year for which the worker flows were calculated.<br>In cases where establishments have closed down, the outflows are shown for the year following the closure. |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.2 Inflows general

#### 5.2.2.1 Inflows (ein_ges)

| Category | Description |
|---|---|
| Variable label | Inflows |

| Category | Description |
|---|---|
| Variable name | ein_ges |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the total number of worker inflows in the respective year.<br>The inflows of a year are defined as the number of employees who were working in the establishment on the reference date of 30.06. of that year but were not working there on the reference date of the previous year. Employees who join the establishment and leave it again between two reference dates are not recorded by this flow concept based on reference dates. |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.2.2 Inflows, female (ein_f)

| Category | Description |
|---|---|
| Variable label | Inflows, female |
| Variable name | ein_f |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of female workers in the respective year.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | The gender is determined at individual level by comparing information from different sources (e.g. social security notifications, jobseeker registration at the BA). The aim is to determine a person's gender consistently across all sources. In a few exceptional cases this is not correct due to errors in data preparation.<br>This variable is a component of the BHP extension file "worker flows", which is only made available to users on application.<br>The number of inflows of male workers is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.2.2.3 Inflows regular workers (ein_reg)

| Category | Description |
|---|---|
| Variable label | Inflows regular workers |
| Variable name | ein_reg |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of inflows of regular workers in the respective year, i.e. individuals who are reported with the person group codes 101, 113, 114, 140 and 143. Further information on the definition of the person groups can be found in Section 8.2.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges) |
| Sensitive variable | No |

| Category | Description |
|---|---|
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.2.4 Inflows marginal part-time (ein_gf)

| Category | Description |
|---|---|
| Variable label | Inflows marginal part-time |
| Variable name | ein_gf |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows into marginal part-time employment in the respective year, i.e. individuals who are reported with the person group codes 109 and 209. Further information on the definition of the person groups can be found in the Section 8.2.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges) |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.2.5 Inflows trainees/apprentices (ein_azubi)

| Category | Description |
|---|---|
| Variable label | Inflows trainees/apprentices |
| Variable name | ein_azubi |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of inflows of trainees/ apprentices in the respective year, i.e. individuals who are identified as trainees/apprentices in the person group code (102, 121, 122, 141, 144). Further information on the person group code can be found in the Appendix (Section 8.2).<br>Detailed information on the definition of inflows can be found in  Inflows (ein_ges) |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.2.6 Inflows full-time (regular workers + others) (ein_vz)

| Category | Description |
|---|---|
| Variable label | Inflows full-time (regular workers + others) |
| Variable name | ein_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of inflows of full-time employees in the respective year, i.e. individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in |

| Category | Description |
|---|---|
| | marginal part-time employment or in partial retirement are not taken into account.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.2.7 Inflows part-time (regular workers + others) (ein_tz)

| Category | Description |
|---|---|
| Variable label | Inflows part-time (regular workers + others) |
| Variable name | ein_tz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of inflows of part-time employees in the respective year, i.e. individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as part-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.2.8 Inflows full-time, female (ein_f_vz)

| Category | Description |
|---|---|
| Variable label | Inflows full-time, female |
| Variable name | ein_f_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of inflows of full-time female employees in the respective year, i.e. women who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

| Category | Description |
|---|---|
| | The number of inflows of full-time male employees is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.2.2.9 Inflows part-time female employees (ein_f_tz)

| Category | Description |
|---|---|
| Variable label | Inflows part-time female employees |
| Variable name | ein_f_tz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of inflows of part-time female employees in the respective year, i.e. women who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as part-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. The number of inflows of part-time male employees is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.2.2.10 Inflows full-time regular workers (ein_reg_vz)

| Category | Description |
|---|---|
| Variable label | Inflows full-time regular workers |
| Variable name | ein_reg_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of inflows of full-time regular workers in the respective year, i.e. individuals who are reported with the person group codes 101, 113, 114, 140 and 143 and as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. Detailed information on the definition of inflows can be found Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.2.11 Inflows, re-hiring (ein_wdr)

| Category | Description |
|---|---|
| Variable label | Inflows, re-hiring |
| Variable name | ein_wdr |
| Data type | Numerical |
| Origin | BeH |

| Category | Description |
|---|---|
| Detailed description | This variable contains the number of inflows in the respective year which are cases of re-hiring.<br><br>A case is regarded as a re-hire if the employee was working in the establishment on the reference date of 30.06 of the respective year and on at least one of the reference dates (30.06) in the preceding 3 years but not on that of the previous year (employment in t, t-2 or t-3, but not in t-1).<br><br>Due to the definition of re-hiring, this variable is only available from 1977 onwards. |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.2.12 Inflows re-hiring, female (ein_wdr_f)

| Category | Description |
|---|---|
| Variable label | Inflows re-hiring, female |
| Variable name | ein_wdr_f |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of female workers in the respective year which are cases of re-hiring.<br><br>A case is regarded as a re-hire if the employee was working in the establishment on the reference date of 30.06 of the respective year and on at least one of the reference dates (30.06) in the preceding 3 years but not on that of the previous year (employment in t, t-2 or t-3, but not in t-1).<br><br>Due to the definition of re-hiring, this variable is only available from 1977 onwards. |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.2.13 Inflows, estab. change (ein_bw)

| Category | Description |
|---|---|
| Variable label | Inflows, estab. change |
| Variable name | ein_bw |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows in the respective year which are cases of establishment movers.<br><br>This means that on 30.06 of the previous year the person was not employed in the current establishment but in a different one.<br><br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.2.14 Inflows estab. change, female (ein_bw_f)

| Category | Description |
|---|---|
| Variable label | Inflows estab. change, female |

| Category | Description |
|---|---|
| Variable name | ein_bw_f |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of female workers in the respective year which are cases of establishment movers.<br>This means that on 30.06 of the previous year the person was not employed in the current establishment but in a different one.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.3 Inflows of employees by Blossfeld occupational group

#### 5.2.3.1 Inflows agricultural occupations (ein_bf_agr)

| Category | Description |
|---|---|
| Variable label | Inflows agricultural occupations |
| Variable name | ein_bf_agr |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of employees in agricultural occupations according to the Blossfeld classification of occupations. Detailed information on the definition of inflows can be found in Inflows (ein_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

#### 5.2.3.2 Inflows unskilled manual occupations (ein_bf_emb)

| Category | Description |
|---|---|
| Variable label | Inflows unskilled manual occupations |
| Variable name | ein_bf_emb |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of employees in unskilled manual occupations according to the Blossfeld classification of occupations. Detailed information on the definition of inflows can be found in Inflows (ein_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable |

| Category | Description |
|---|---|
| | has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.3.3  Inflows unskilled services (ein_bf_edi)

| Category | Description |
|---|---|
| Variable label | Inflows unskilled services |
| Variable name | ein_bf_edi |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of employees in unskilled services occupations according to the Blossfeld classification of occupations. Detailed information on the definition of inflows can be found in Inflows (ein_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.3.4  Inflows unskilled commercial and admin. occupations (ein_bf_evb)

| Category | Description |
|---|---|
| Variable label | Inflows unskilled commercial and admin. occupations |
| Variable name | ein_bf_evb |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of employees in unskilled commercial and administrative occupations according to the Blossfeld classification of occupations. Detailed information on the definition of inflows can be found in Inflows (ein_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.3.5  Inflows skilled manual occupations (ein_bf_qmb)

| Category | Description |
|---|---|
| Variable label | Inflows skilled manual occupations |
| Variable name | ein_bf_qmb |
| Data type | Numerical |

| Category | Description |
|---|---|
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of employees in skilled manual occupations according to the Blossfeld classification of occupations. <br><br> Detailed information on the definition of inflows can be found in Inflows (ein_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.3.6  Inflows skilled services (ein_bf_qdi)

| Category | Description |
|---|---|
| Variable label | Inflows skilled services |
| Variable name | ein_bf_qdi |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of employees in skilled manual services according to the Blossfeld classification of occupations. Detailed information on the definition of inflows can be found in Inflows (ein_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.3.7  Inflows skilled commercial and admin. occupations (ein_bf_qvb)

| Category | Description |
|---|---|
| Variable label | Inflows skilled commercial and admin. occupations |
| Variable name | ein_bf_qvb |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of employees in skilled commercial and administrative occupations according to the Blossfeld classification of occupations. <br><br> Detailed information on the definition of inflows can be found in Inflows (ein_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable |

| Category | Description |
|---|---|
| | has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.3.8 Inflows technicians (ein_bf_tec)

| Category | Description |
|---|---|
| Variable label | Inflows technicians |
| Variable name | ein_bf_tec |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of employed technicians according to the Blossfeld classification of occupations. Detailed information on the definition of inflows can be found in Inflows (ein_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.3.9 Inflows semiprofessions (ein_bf_semi)

| Category | Description |
|---|---|
| Variable label | Inflows semiprofessions |
| Variable name | ein_bf_semi |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of employed semiprofessions according to the Blossfeld classification of occupations. Detailed information on the definition of inflows can be found in Inflows (ein_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.3.10 Inflows engineers (ein_bf_ing)

| Category | Description |
|---|---|
| Variable label | Inflows engineers |
| Variable name | ein_bf_ing |
| Data type | Numerical |
| Origin | BeH |

| Category | Description |
|---|---|
| Detailed description | This variable contains the number of inflows of employed engineers according to the Blossfeld classification of occupations.<br><br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.3.11 Inflows professions (ein_bf_prof)

| Category | Description |
|---|---|
| Variable label | Inflows professions |
| Variable name | ein_bf_prof |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of employed professions according to the Blossfeld classification of occupations.<br><br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.3.12 Inflows managers (ein_bf_man)

| Category | Description |
|---|---|
| Variable label | Inflows manager |
| Variable name | ein_bf_man |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of inflows of employed managers according to the Blossfeld classification of occupations.<br><br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). |

| Category | Description |
|---|---|
| | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.4 Inflows by age class

#### 5.2.4.1 Inflows aged 15–19 (ein_15_19)

| Category | Description |
|---|---|
| Variable label | Inflows aged 15-19 |
| Variable name | ein_15_19 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows aged between 15 and 19 in the respective year. Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

#### 5.2.4.2 Inflows aged 20–24 (ein_20_24)

| Category | Description |
|---|---|
| Variable label | Inflows aged 20-24 |
| Variable name | ein_20_24 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows aged between 20 and 24 in the respective year. Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

#### 5.2.4.3 Inflows aged 25–29 (ein_25_29)

| Category | Description |
|---|---|
| Variable label | Inflows aged 25-29 |
| Variable name | ein_25_29 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows aged between 25 and 29 in the respective year. Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

#### 5.2.4.4 Inflows aged 30–34 (ein_30_34)

| Category | Description |
|---|---|
| Variable label | Inflows aged 30-34 |

| Category | Description |
|---|---|
| Variable name | ein_30_34 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows aged between 30 and 34 in the respective year.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.4.5  Inflows aged 35–39 (ein_35_39)

| Category | Description |
|---|---|
| Variable label | Inflows aged 35-39 |
| Variable name | ein_35_39 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows aged between 35 and 39 in the respective year.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.4.6  Inflows aged 40–44 (ein_40_44)

| Category | Description |
|---|---|
| Variable label | Inflows aged 40-44 |
| Variable name | ein_40_44 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows aged between 40 and 44 in the respective year.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.4.7  Inflows aged 45–49 (ein_45_49)

| Category | Description |
|---|---|
| Variable label | Inflows aged 45-49 |
| Variable name | ein_45_49 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows aged between 45 and 49 in the respective year.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |

| Category | Description |
|---|---|
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.4.8 Inflows aged 50–54 (ein_50_54)

| Category | Description |
|---|---|
| Variable label | Inflows aged 50-54 |
| Variable name | ein_50_54 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows aged between 50 and 54 in the respective year.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.4.9 Inflows aged 55–59 (ein_55_59)

| Category | Description |
|---|---|
| Variable label | Inflows aged 55-59 |
| Variable name | ein_55_59 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows aged between 55 and 59 in the respective year.<br>Detailed information on the definition of inflows can be found in Section 5.2.2.1. |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.4.10 Inflows aged 60–64 (ein_60_64)

| Category | Description |
|---|---|
| Variable label | Inflows aged 60-64 |
| Variable name | ein_60_64 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows aged between 60 and 64 in the respective year.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.4.11 Inflows aged >64 (ein_ab65)

| Category | Description |
|---|---|
| Variable label | Inflows aged >64 |
| Variable name | ein_ab65 |

| Category | Description |
|---|---|
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker inflows aged 65 or above in the respective year.<br>Detailed information on the definition of inflows can be found in Inflows (ein_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

## 5.2.5 Outflows, general

### 5.2.5.1 Outflows (aus_ges)

| Category | Description |
|---|---|
| Variable label | Outflows |
| Variable name | aus_ges |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the total number of worker outflows in the respective year.<br>The outflows of a year are defined as the number of employees who were not working in the establishment on the reference date of 30.06 of that year but were working there on the reference date of the previous year. Employees who join the establishment and leave it again between two reference dates are not recorded by this flow concept based on reference dates. |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.5.2 Outflows, female (aus_f)

| Category | Description |
|---|---|
| Variable label | Outflows, female |
| Variable name | aus_f |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of female workers in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year.<br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | The gender is determined at individual level by comparing information from different sources (e.g. social security notifications, jobseeker registration at the BA). The aim is to determine a person's gender consistently across all sources. In a few exceptional cases this is not correct due to errors in data preparation. |

| Category | Description |
|---|---|
| | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |
| | The number of outflows of male workers is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.2.5.3  Outflows regular workers (aus_reg)

| Category | Description |
|---|---|
| Variable label | Outflows regular workers |
| Variable name | aus_reg |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of outflows of regular workers in the respective year, i.e. individuals who are reported with the person group codes 101, 113, 114, 140 and 143. Further information on the definition of the person groups can be found in the Section 8.2. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.5.4  Outflows marginal part-time (aus_gf)

| Category | Description |
|---|---|
| Variable label | Outflows marginal part-time |
| Variable name | aus_gf |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows from marginal part-time employment in the respective year, i.e. individuals who are reported with the person group codes 109 and 209. Further information on the definition of the person groups can be found in the Section 8.2. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.5.5  Outflows trainees/apprentices (aus_azubi)

| Category | Description |
|---|---|
| Variable label | Outflows trainees/apprentices |
| Variable name | aus_azubi |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of outflows of trainees/ apprentices in the respective year, i.e. individuals who are identified as |

| Category | Description |
|---|---|
|  | trainees/apprentices in the person group code (102, 121, 122, 141, 144). Further information on the person group code can be found in the Appendix (Section 8.2). |
|  | It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. |
|  | Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.5.6  Outflows full-time (regular workers + others) (aus_vz)

| Category | Description |
|---|---|
| Variable label | Outflows full-time (regular workers + others) |
| Variable name | aus_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of outflows of full-time employees in the respective year, i.e. individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
|  | It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. |
|  | Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.5.7  Outflows part-time (regular workers + others) (aus_tz)

| Category | Description |
|---|---|
| Variable label | Outflows part-time (regular workers + others) |
| Variable name | aus_tz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of outflows of part-time employees in the respective year, i.e. individuals who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as part-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in |

| Category | Description |
|---|---|
| | marginal part-time employment or in partial retirement are not taken into account. |
| | It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. |
| | Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.5.8 Outflows full-time, female (aus_f_vz)

| Category | Description |
|---|---|
| Variable label | Outflows full-time, female |
| Variable name | aus_f_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of outflows of full-time female employees in the respective year, i.e. women who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as full-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and people in marginal part-time employment or in partial retirement are not taken into account. |
| | It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. |
| | Detailed information on the definition of outflows can be found in v. |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |
| | The number of outflows of full-time male employees is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.2.5.9 Outflows part-time female employees (aus_f_tz)

| Category | Description |
|---|---|
| Variable label | Outflows part-time female employees |
| Variable name | aus_f_tz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of outflows of part-time female employees in the respective year, i.e. women who are reported with the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 and YYY and as part-time employees. Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6. This means that trainees/apprentices and |

| Category | Description |
|---|---|
| | people in marginal part-time employment or in partial retirement are not taken into account.<br><br>It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year.<br><br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application.<br><br>The number of outflows of part-time male employees is not included in the BHP but can be calculated by the user (see Section 3.10). |

### 5.2.5.10 Outflows full-time regular workers (aus_reg_vz)

| Category | Description |
|---|---|
| Variable label | Outflows full-time regular workers |
| Variable name | aus_reg_vz |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | The variable contains the number of outflows of full-time regular workers in the respective year, i.e. individuals who are reported with the person group codes 101, 113, 114, 140 and 143 and as full-time employees.<br><br>Further information on the person group code can be found in the Appendix (Section 8.2) and information on the definition of full-time and part-time employment in Section 3.1.6.<br><br>It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year.<br><br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.5.11 Outflows temp. (aus_temp)

| Category | Description |
|---|---|
| Variable label | Outflows temp. |
| Variable name | aus_temp |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows in the respective year that are temporary outflows.<br><br>An outflow is regarded as temporary if the employee was not working in the establishment on the reference date of the 30.06. of the respective year but was working there on the reference date of the previous year and on the reference date of the following year (employment in t-1 and t+1 but not in t). |
| Sensitive variable | No |

| Category | Description |
|---|---|
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.5.12 Outflows temp., female (aus_temp_f)

| Category | Description |
|---|---|
| Variable label | Outflow temp., female |
| Variable name | aus_temp_f |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of female workers in the respective year that are temporary outflows.<br>An outflow is regarded as temporary if the employee was not working in the establishment on the reference date of the 30.06. of the respective year but was working there on the reference date of the previous year and on the reference date of the following year (employment in t-1 and t+1 but not in t). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.5.13 Outflows, estab. change (aus_bw)

| Category | Description |
|---|---|
| Variable label | Outflows, estab. change |
| Variable name | aus_bw |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows in the respective year that move to a different establishment. This means that the person was employed in the establishment on 30.06 of the previous year but is currently employed in a different establishment.<br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.5.14 Outflows estab. change, female (aus_bw_f)

| Category | Description |
|---|---|
| Variable label | Outflows estab. change, female |
| Variable name | aus_bw_f |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of female workers in the respective year that move to a different establishment. This means that the person was employed in the establishment on 30.06 of the previous year but is currently employed in a different establishment.<br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |

| Category | Description |
|---|---|
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

## 5.2.6 Outflows of employees by Blossfeld occupational group

### 5.2.6.1 Outflows agricultural occupations (aus_bf_agr)

| Category | Description |
|---|---|
| Variable label | Outflows agricultural occupations |
| Variable name | aus_bf_agr |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of employees in agricultural occupations according to the Blossfeld classification of occupations in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. Detailed information on the definition of outflows can be found in Outflows (aus_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.6.2 Outflows unskilled manual occupations (aus_bf_emb)

| Category | Description |
|---|---|
| Variable label | Outflows unskilled manual occupations |
| Variable name | aus_bf_emb |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of employees in unskilled manual occupations according to the Blossfeld classification of occupations in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. Detailed information on the definition of outflows can be found in Outflows (aus_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.6.3 Outflows unskilled services (aus_bf_edi)

| Category | Description |
|---|---|
| Variable label | Outflows unskilled services |
| Variable name | aus_bf_edi |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of employees in unskilled services according to the Blossfeld classification of occupations in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year.<br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.6.4 Outflows unskilled commercial and admin. occupations (aus_bf_evb)

| Category | Description |
|---|---|
| Variable label | Outflows unskilled commercial and admin. occupations |
| Variable name | aus_bf_evb |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of employees in unskilled commercial and administrative occupations according to the Blossfeld classification of occupations in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year.<br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.6.5 Outflows skilled manual occupations (aus_bf_qmb)

| Category | Description |
|---|---|
| Variable label | Outflows skilled manual occupations |
| Variable name | aus_bf_qmb |
| Data type | Numerical |
| Origin | BeH |

| Category | Description |
|---|---|
| Detailed description | This variable contains the number of outflows of employees in skilled manual occupations according to the Blossfeld classification of occupations in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year.<br><br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6).<br><br>This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.6.6 Outflows skilled services (aus_bf_qdi)

| Category | Description |
|---|---|
| Variable label | Outflows skilled services |
| Variable name | aus_bf_qdi |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable contains the number of outflows of employees in skilled services occupations according to the Blossfeld classification of occupations in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year.<br><br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.6.7 Outflows skilled commercial and admin. occupations (aus_bf_qvb)

| Category | Description |
|---|---|
| Variable label | Outflows skilled commercial and admin. occupations |
| Variable name | aus_bf_qvb |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of employees in skilled commercial and administrative occupations according to the Blossfeld classification of occupations in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. |

| Category | Description |
|---|---|
| | Detailed information on the definition of outflows can be found in Outflows (aus_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.6.8 Outflows technicians (aus_bf_tec)

| Category | Description |
|---|---|
| Variable label | Outflows technicians |
| Variable name | aus_bf_tec |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of employed technicians according to the Blossfeld classification of occupations in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. Detailed information on the definition of outflows can be found in Outflows (aus_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.6.9 Outflows semiprofessions (aus_bf_semi)

| Category | Description |
|---|---|
| Variable label | Outflows semiprofessions |
| Variable name | aus_bf_semi |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of employed semiprofessions according to the Blossfeld classification of occupations in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. Detailed information on the definition of outflows can be found in Outflows (aus_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |

| Category | Description |
|---|---|
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.6.10 Outflows engineers (aus_bf_ing)

| Category | Description |
|---|---|
| Variable label | Outflows engineers |
| Variable name | aus_bf_ing |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of employed engineers according to the Blossfeld classification of occupations in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year.<br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.6.11 Outflows professions (aus_bf_prof)

| Category | Description |
|---|---|
| Variable label | Outflows professions |
| Variable name | aus_bf_prof |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of employed professions according to the Blossfeld classification of occupations in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year.<br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.6.12 Outflows managers (aus_bf_man)

| Category | Description |
|---|---|
| Variable label | Outflows managers |
| Variable name | aus_bf_man |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of outflows of employed managers according to the Blossfeld classification of occupations in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. |
| | Detailed information on the definition of outflows can be found in Outflows (aus_ges). Detailed information on the Blossfeld classification of occupations can be found in No. agricultural occupations (az_bf_agr). |
| Sensitive variable | No |
| Notes | Due to the changeover from the Classification of Occupations 1988 (KldB 1988) to the Classification of Occupations 2010 (KldB 2010) the variable has significant gaps for 2011. From 2012 onwards, the values of the KldB 2010 were recoded to the categories of the KldB 1988 (see Section 4.6). This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

## 5.2.7 Outflows by age class

### 5.2.7.1 Outflows aged 15–19 (aus_15_19)

| Category | Description |
|---|---|
| Variable label | Outflows aged 15-19 |
| Variable name | aus_15_19 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows aged between 15 and 19 in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. |
| | Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.7.2 Outflows aged 20–24 (aus_20_24)

| Category | Description |
|---|---|
| Variable label | Outflows aged 20-24 |
| Variable name | aus_20_24 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows aged between 20 and 24 in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. |

| Category | Description |
|---|---|
| | Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.7.3 Outflows aged 25–29 (aus_25_29)

| Category | Description |
|---|---|
| Variable label | Outflows aged 25-29 |
| Variable name | aus_25_29 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows aged between 25 and 29 in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.7.4 Outflows aged 30–34 (aus_30_34)

| Category | Description |
|---|---|
| Variable label | Outflows aged 30-34 |
| Variable name | aus_30_34 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows aged between 30 and 34 in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.7.5 Outflows aged 35–39 (aus_35_39)

| Category | Description |
|---|---|
| Variable label | Outflows aged 35-39 |
| Variable name | aus_35_39 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows aged between 35 and 39 in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. |

| Category | Description |
|---|---|
|  | Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.7.6 Outflows aged 40-44 (aus_40_44)

| Category | Description |
|---|---|
| Variable label | Outflows aged 40-44 |
| Variable name | aus_40_44 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows aged between 40 and 44 in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.7.7 Outflows aged 45–49 (aus_45_49)

| Category | Description |
|---|---|
| Variable label | Outflows aged 45-49 |
| Variable name | aus_45_49 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows aged between 45 and 49 in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.7.8 Outflows aged 50–54 (aus_50_54)

| Category | Description |
|---|---|
| Variable label | Outflows aged 50-54 |
| Variable name | aus_50_54 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows aged between 50 and 54 in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. |

| Category | Description |
|---|---|
| | Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.7.9 Outflows aged 55–59 (aus_55_59)

| Category | Description |
|---|---|
| Variable label | Outflows aged 55-59 |
| Variable name | aus_55_59 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows aged between 55 and 59 in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year.<br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.7.10 Outflows aged 60–64 (aus_60_64)

| Category | Description |
|---|---|
| Variable label | Outflows aged 60-64 |
| Variable name | aus_60_64 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows aged between 60 and 64 in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year.<br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.7.11 Outflows aged >64 (aus_ab65)

| Category | Description |
|---|---|
| Variable label | outflows aged >64 |
| Variable name | aus_ab65 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows aged 65 or above in the respective year. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. |

| Category | Description |
|---|---|
| | Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

### 5.2.8 Job tenure

#### 5.2.8.1 Outflows job tenure < 4 years (aus_senio_1)

| Category | Description |
|---|---|
| Variable label | Outflows job tenure <4 years |
| Variable name | aus_senio_1 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows in the respective year after 1-3 years of job tenure. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

#### 5.2.8.2 Outflows job tenure 4–9 years (aus_senio_2)

| Category | Description |
|---|---|
| Variable label | Outflows job tenure 4-9 years |
| f | aus_senio_2 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows in the respective year after 4-9 years of job tenure. It must be taken into account that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year. Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

#### 5.2.8.3 Outflows job tenure > 9 years (aus_senio_3)

| Category | Description |
|---|---|
| Variable label | Outflows job tenure >9 years |
| Variable name | aus_senio_3 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of worker outflows in the respective year after 10 or more years of job tenure. It must be taken into account |

| Category | Description |
|---|---|
| | that the value 0 can also be attributed to the employee group not being represented in the establishment in the previous year.<br>Detailed information on the definition of outflows can be found in Outflows (aus_ges). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "worker flows", which is only made available to users on application. |

## 5.3    Extension file – entry and exit

### 5.3.1    Establishment variables

#### 5.3.1.1  Establishment ID (bhp_betnr)

| Category | Description |
|---|---|
| Variable label | Establishment id |
| Variable name | bhp_betnr |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | To make the data anonymous, the original establishment numbers allocated by the Federal Employment Agency are replaced by randomly generated, yet unambiguous establishment IDs. Further information on the allocation of establishment numbers and on the definition of an establishment can be found in Section 2.2.<br>The artificial establishment ID may be used to merge the yearly data of the core dataset to create a panel dataset and to add the extension files and the file with the time-consistent classifications of economic activities to the sets of yearly data of the core dataset. |
| Sensitive variable | No |

#### 5.3.1.2  Year (jahr)

| Category | Description |
|---|---|
| Variable label | Year |
| Variable name | Jahr |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the year for which the entries and exits were calculated. In the case of establishment closures, the exits are shown for the year following the closure. |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "entry and exit", which is only made available to users on application. |

### 5.3.2    Entry

#### 5.3.2.1  Entry type (eintritt)

| Category | Description |
|---|---|
| Variable label | Entry type |
| Variable name | eintritt |
| Data type | Numerical |

| Category | Description |
| --- | --- |
| Origin | BeH |
| Detailed description | This variable contains the establishment's entry status. A distinction is made between whether it is a genuine new establishment start-up or only a spin-off of part of an existing establishment or a change of establishment number.<br>The entry status is assigned by examining worker flows in the year of entry. Detailed information about this procedure can be found in Section 3.5.1 and in Hethey and Schmieder (2010).<br>Variable label:<br>1   ID change<br>2   Spin-off/pulled<br>3   Spin-off/pushed<br>4   New estab. (small)<br>5   New estab. (medium & large)<br>6   New estab. (chunky)<br>7   Unclear |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "entry and exit", which is only made available to users on application. |

### 5.3.2.2  Employment of bhp_betnr in year (besch)

| Category | Description |
| --- | --- |
| Variable label | Employment of betnr in year |
| Variable name | besch |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the total number of employees as of the reference date of the year of entry.<br>The variable is an auxiliary variable to determine an establishment's entry status (see the variable "Entry type (eintritt)"). Detailed information about this procedure can be found in Section 3.5.1 and in Hethey and Schmieder (2010). |
| Sensitive variable | No |
| Notes | This variable is a component of the BHP extension file "entry and exit", which is only made available to users on application. |

### 5.3.2.3  Inflow from predecessor to bhp_betnr (inflow)

| Category | Description |
| --- | --- |
| Variable label | Inflow from predecessor to betnr |
| Variable name | inflow |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the number of employees who switch from the predecessor establishment to the establishment under examination (as of the reference date of the year of entry). Only the predecessor establishment from which the most employees switch to the new establishment – the maximum clustered inflow – is taken into account. |

| Category | Description |
|---|---|
| | The variable is an auxiliary variable to determine an establishment's entry status (see the variable "Entry type (eintritt)"). Detailed information about this procedure can be found in Section 3.5.1 and in Hethey and Schmieder (2010). |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "entry and exit", which is only made available to users on application. |

### 5.3.2.4 Employment predecessor one year before (besch_vor)

| Category | Description |
|---|---|
| Variable label | Employment predecessor one year before |
| Variable name | besch_vor |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the total number of employees in the predecessor establishment as of the reference date the year before the establishment was founded<br>The variable is an auxiliary variable to determine an establishment's entry status (see the variable "Entry type (eintritt)"). Detailed information about this procedure can be found in Section 3.5.1 and in Hethey and Schmieder (2010). |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "entry and exit", which is only made available to users on application. |

### 5.3.2.5 Predecessor exits (status_vor)

| Category | Description |
|---|---|
| Variable label | Predecessor exits |
| Variable name | status_vor |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains information about the status of the predecessor establishment. It distinguishes whether the establishment still exists or is being closed down.<br>The variable is an auxiliary variable to determine an establishment's entry status (see the variable "Entry type (eintritt)"). Detailed information about this procedure can be found in Section 3.5.1 and in Hethey and Schmieder (2010).<br>Values of the variable:<br>    0    Predecessor continues<br>    1    Predecessor exits |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "entry and exit", which is only made available to users on application. |

### 5.3.3 Exit

#### 5.3.3.1 Exit type (austritt)

| Category | Description |
|---|---|
| Variable label | Exit type |
| Variable name | austritt |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains details regarding the type of exit of the establishment. It distinguishes first and foremost whether it is a genuine closure or simply a take-over by another establishment or a change of establishment number.<br>The exit status is assigned by examining worker flows in the year of exit. Detailed information about this procedure can be found in Section 3.5.1 and in Hethey and Schmieder (2010).<br>Variable label:<br>    1   ID change<br>    2   Take-over/restructuring<br>    3   Spin-off/pushed<br>    4   Small death<br>    5   Atomized death<br>    6   Chunky death<br>    7   Unclear |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "entry and exit", which is only made available to users on application. |

#### 5.3.3.2 Employment of betnr in year (besch)

| Category | Description |
|---|---|
| Variable label | Employment of betnr in year |
| Variable name | besch |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the total number of employees as of the reference date of the year of exit.<br>The variable is an auxiliary variable to determine an establishment's exit status (see the variable "Exit type (austritt)"). Detailed information about this procedure can be found in Section 3.5.1 and in Hethey and Schmieder (2010). |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "entry and exit, which is only made available to users on application. |

#### 5.3.3.3 Outflow from betnr to successor (outflow)

| Category | Description |
|---|---|
| Variable label | Outflow from betnr to successor |
| Variable name | outflow |
| Data type | Numerical |
| Origin | BeH |

| Category | Description |
|---|---|
| Detailed description | This variable contains the number of employees who switch to the successor establishment after the exit of the establishment under examination. Only the successor establishment to which the most employees switch from the old establishment (the maximum clustered outflow) is taken into account.<br><br>The variable is an auxiliary variable to determine an establishment's exit status (see the variable "Entry type (eintritt)"). Detailed information about this procedure can be found in Section 3.5.1 and in Hethey and Schmieder (2010). |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "entry and exit", which is only made available to users on application. |

### 5.3.3.4  Employment successor one year later (besch_nach)

| Category | Description |
|---|---|
| Variable label | Employment successor one year later |
| Variable name | besch_nach |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the total number of employees in the successor establishment as of the reference date of the year following the closure.<br><br>The variable is an auxiliary variable to determine an establishment's exit status (see the variable "Entry type (eintritt)"). Detailed information about this procedure can be found in Section 3.5.1 and in Hethey and Schmieder (2010). |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "entry and exit", which is only made available to users on application. |

### 5.3.3.5  Successor is entrant (status_nach)

| Category | Description |
|---|---|
| Variable label | Successor is entrant |
| Variable name | status_nach |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains information about the status of the successor establishment. It distinguishes whether the successor is a new establishment or one that was established previously.<br><br>The variable is an auxiliary variable to determine an establishment's exit status (see the variable "Entry type (eintritt)"). Detailed information about this procedure can be found in Section 3.5.1 and in Hethey and Schmieder (2010).<br><br>Variable label:<br>0    Successor exists<br>1    Successor is entrant |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "entry and exit", which is only made available to users on application. |

## 5.4 Time-consistent industry codes

### 5.4.1 Establishment variables

#### 5.4.1.1 Establishment ID (bhp_betnr)

| Category | Description |
|---|---|
| Variable label | Establishment id |
| Variable name | bhp_betnr |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | To make the data anonymous, the original establishment numbers allocated by the Federal Employment Agency are replaced by randomly generated, yet unambiguous establishment IDs. Further information on the allocation of establishment numbers and on the definition of an establishment can be found in Section 2.2.<br>The artificial establishment ID may be used to merge the yearly data of the core dataset to create a panel dataset and to add the extension files and the file with the time-consistent classifications of economic activities to the sets of yearly data of the core dataset. |
| Sensitive variable | no |

#### 5.4.1.2 Year (jahr)

| Category | Description |
|---|---|
| Variable label | Year |
| Variable name | jahr |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the year for which the time-consistent classifications of economic activities were generated. |
| Sensitive variable | no |

### 5.4.2 Generated variables

#### 5.4.2.1 w73_3 completed by extrapolation/imputation (w73_3_gen)

| Category | Description |
|---|---|
| Variable label | w73_3 completed by extrapolation/imputation |
| Variable name | w73_3_gen |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the completed economic activities in accordance with the classification WS73 as a 3-digit code. A detailed description of this procedure can be found in Eberle et al. (2011). |
| Sensitive variable | no |

#### 5.4.2.2 Type of imputation w73_3 (group_w73_3)

| Category | Description |
|---|---|
| Variable label | Type of imputation w73_3 |
| Variable name | group_w73_3 |
| Data type | Numerical |
| Origin | BeH |

| Category | Description |
|---|---|
| Detailed description | This variable contains the type of imputation used for the variable "w73_3_gen". A detailed description of this procedure can be found in Eberle et al. (2011).<br>Values of the variable:<br>    0   Original value<br>    1   Remains missing<br>    2   Extrapolated<br>    3   Imp. using w93_5 |
| Sensitive variable | no |

### 5.4.2.3   w93_3 completed by extrapolation/imputation (w93_3_gen)

| Category | Description |
|---|---|
| Variable label | w93_3 completed by extrapolation/imputation |
| Variable name | w93_3_gen |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the completed economic activities in accordance with the WZ93 classification as a 3-digit code. A detailed description of this procedure can be found in Eberle et al. (2011). |
| Sensitive variable | no |

### 5.4.2.4   Type of imputation w93_3 (group_w93_3)

| Category | Description |
|---|---|
| Variable label | Type of imputation w93_3 |
| Variable name | group_w93_3 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the type of imputation used for the variable "w93_3_gen". A detailed description of this procedure can be found in Eberle et al. (2011).<br>Values of the variable:<br>    0 Original value<br>    1 Remains missing<br>    2 Extrapolated<br>    3 Imp. using w03_5<br>    4 2nd imp. using w73_3 |
| Sensitive variable | no |

### 5.4.2.5   w08_3 completed by extrapolation/imputation (w08_3_gen)

| Category | Description |
|---|---|
| Variable label | w08_3 completed by extrapolation/imputation |
| Variable name | w08_3_gen |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the completed economic activities in accordance with the WZ08 classification as a 3-digit code. The variable is computed following the method for generating time-consistent industry codes described in Eberle et al. (2011). |
| Sensitive variable | no |

### 5.4.2.6 Type of imputation w08_3 (group_w08_3)

| Category | Description |
|---|---|
| Variable label | Type of imputation w08_3 |
| Variable name | group_w08_3 |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the type of imputation used for the variable "w08_3_gen".<br>Values of the variable:<br>0 Original value<br>1 Remains missing<br>2 Extrapolated<br>3 Imputed |
| Sensitive variable | no |

## 5.5    Extension file – bankruptcies

### 5.5.1    Establishment variables

#### 5.5.1.1 Establishment ID (bhp_betnr)

| Category | Description |
|---|---|
| Variable label | Establishment id |
| Variable name | bhp_betnr |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | To make the data anonymous, the original establishment numbers allocated by the Federal Employment Agency are replaced by randomly generated, yet unambiguous establishment IDs. Further information on the allocation of establishment numbers and on the definition of an establishment can be found in Section 2.2.<br>The artificial establishment ID may be used to merge the yearly data of the core dataset to create a panel dataset and to add the extension files and the file with the time-consistent classifications of economic activities to the sets of yearly data of the core dataset. |
| Sensitive variable | no |

### 5.5.2    Bankruptcies

#### 5.5.2.1 Source: monitoring (iv_monitoring)

| Category | Description |
|---|---|
| Variable label | Source: monitoring |
| Variable name | iv_monitoring |
| Data type | Numerical |
| Origin | Data of the statistics department of the Federal Employment Agency |
| Detailed description | This variable indicates if the "monitoring" data (see section 3.6) was used to identify the bankrupt establishment.<br>Variable label:<br>0    No<br>1    Yes |
| Sensitive variable | no |

| Category | Description |
|---|---|
| Notes | This variable is a component of the BHP extension file "bankruptcies", which is only made available to users on application. |

### 5.5.2.2 Source: employment notification (iv_meldung)

| Category | Description |
|---|---|
| Variable label | Source: employment notification |
| Variable name | iv_meldung |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable indicates if the "notification" data (see section 3.6) was used to identify the bankrupt establishment.<br>Variable label:<br>    0    No<br>    1    Yes |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "bankruptcies", which is only made available to users on application. |

### 5.5.2.3 Source: application for Insolvenzgeld (iv_geld)

| Category | Description |
|---|---|
| Variable label | Source: application for Insolvenzgeld |
| Variable name | iv_geld |
| Data type | Numerical |
| Origin | Data of the statistics department of the Federal Employment Agency |
| Detailed description | This variable indicates if the "application for Insolvenzgeld" data (see section 3.6) was used to identify the bankrupt establishment.<br>Variable label:<br>    0    No<br>    1    Yes |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "bankruptcies", which is only made available to users on application. |

### 5.5.2.4 First year: monitoring (iv_monitoring_begjahr)

| Category | Description |
|---|---|
| Variable label | First year: monitoring |
| Variable name | iv_monitoring_begjahr |
| Data type | Numerical |
| Origin | Data of the statistics department of the Federal Employment Agency |
| Detailed description | This variable contains the first year of the time span the "monitoring" data (see section 3.6) on bankruptcies refer to. |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "bankruptcies", which is only made available to users on application. |

### 5.5.2.5 First year: employment notification (iv_meldung_begjahr)

| Category | Description |
|---|---|
| Variable label | First year: employment notification |
| Variable name | iv_meldung_begjahr |
| Data type | Numerical |

| Category | Description |
|---|---|
| Origin | BeH |
| Detailed description | This variable contains the first year of the time span the "notification" data (see section 3.6) on bankruptcies refer to. |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "bankruptcies", which is only made available to users on application. |

### 5.5.2.6 Last year: employment notification (iv_meldung_endjahr)

| Category | Description |
|---|---|
| Variable label | Last year: employment notification |
| Variable name | iv_meldung_endjahr |
| Data type | Numerical |
| Origin | BeH |
| Detailed description | This variable contains the last year of the time span the "notification" data (see section 3.6) on bankruptcies refer to |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "bankruptcies", which is only made available to users on application. |

### 5.5.2.7 First year: application for Insolvenzgeld (iv_geld_begjahr)

| Category | Description |
|---|---|
| Variable label | First year: application for Insolvenzgeld |
| Variable name | iv_geld_begjahr |
| Data type | Numerical |
| Origin | Data of the statistics department of the Federal Employment Agency |
| Detailed description | This variable contains the first year of the time span the "application for Insolvenzgeld" data (see section 3.6) on bankruptcies refer to. |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "bankruptcies", which is only made available to users on application. |

### 5.5.2.8 Last year: application for Insolvenzgeld (iv_geld_endjahr)

| Category | Description |
|---|---|
| Variable label | Last year: application for Insolvenzgeld |
| Variable name | iv_geld_endjahr |
| Data type | Numerical |
| Origin | Data of the statistics department of the Federal Employment Agency |
| Detailed description | This variable contains the last year of the time span the "application for Insolvenzgeld" data (see section 3.6) on bankruptcies refer to. |
| Sensitive variable | no |
| Notes | This variable is a component of the BHP extension file "bankruptcies", which is only made available to users on application. |

# 6 References

**Bender, Stefan; Hilzendegen, Jürgen; Rohwer, Götz; Rudolph, Helmut** (1996): Die IAB-Beschäftigtenstichprobe 1975-1990. Beiträge zur Arbeitsmarkt- und Berufsforschung 197, Nürnberg (only available in German).

**Bertat, Thomas; Dundler, Agnes; Grimm, Christopher; Kiewitt, Jochen; Schomaker, Christine; Schridde, Henning; Zemann, Christian** (2013): Neue Erhebungsinhalte „Arbeitszeit", „ausgeübte Tätigkeit", sowie „Schul- und Berufsabschluss" in der Beschäftigungsstatistik. Methodenbericht der Statistik der Bundesagentur für Arbeit, Nürnberg. URL: https://statistik.arbeitsagentur.de/DE/Statischer-Content/Grundlagen/Methodik-Qualitaet/Methodenberichte/Beschaeftigungsstatistik/Generische-Publikationen/MB-Arbeitszeit-ausgeuebte-Taetigkeit-Schul-und-Berufsabschluss.pdf?__blob=publicationFile&v=4 (Accessed: 30.11.2022, only available in German).

**Blossfeld, Hans-Peter** (1987): Labor Market Entry and the Sexual Segregation of Careers in the Federal Republic of Germany. In: The American Journal of Sociology 93 (1), pp. 89–118.

**Bundesagentur für Arbeit (Ed.)** (2007): Handbuch für die Betriebsnummernvergabe und -pflege im Rahmen des Meldeverfahrens zur Sozialversicherung. Edition December 2007, Nürnberg (only available in German).

**Bundesagentur für Arbeit (Hg.)** (2021a): Klassifikation der Berufe 2010 – überarbeitete Fassung 2020. Band 1: Systematischer und alphabetischer Teil mit Erläuterungen, Nürnberg. URL: https://statistik.arbeitsagentur.de/DE/Statischer-Content/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB2010-Fassung2020/Printausgabe-KldB-2010-Fassung2020/Generische-Publikationen/KldB2010-PDF-Version-Band1-Fassung2020.pdf?__blob=publicationFile&v=21 (Accessed: 30.11.2022, only available in German).

**Bundesagentur für Arbeit (Hg.) (2021b):** Klassifikation der Berufe 2010 – überarbeitete Fassung 2020. Band 2: Definitorischer und beschreibender Teil, Nürnberg. URL: https://statistik.arbeitsagentur.de/DE/Statischer-Content/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB2010-Fassung2020/Printausgabe-KldB-2010-Fassung2020/Generische-Publikationen/KldB2010-PDF-Version-Band2-Fassung2020.pdf?__blob=publicationFile&v=20 (Accessed: 30.11.2022, only available in German).

**Bundesanstalt für Arbeit (Ed.)** (1988): Klassifizierung der Berufe – Systematisches und Alphabetisches Verzeichnis der Berufsbenennungen, Nürnberg. URL: https://statistik.arbeitsagentur.de/DE/Statischer-Content/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB1975-1992/Generische-Publikationen/KldB1988-Systematisches-und-alphabetisches-Verzeichnis-der-Berufsbenennungen-PDF.pdf?__blob=publicationFile&v=12 (Accessed: 30.11.2022, only available in German).

**Bundesanstalt für Arbeit (Ed.)** (1996): Klassifikation der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit – WZ93/BA, Nürnberg (only available in German).

**Card, David; Heining, Jörg; Kline, Patrick** (2015): CHK effects. FDZ-Methodenreport, 06/2015 (en), Nürnberg.

**Cramer, Ulrich** (1985): Probleme der Genauigkeit der Beschäftigtenstatistik. In: Allgemeines Statistisches Archiv 69, pp. 56–68 (only available in German).

**Eberle, Johanna; Jacobebbinghaus, Peter; Ludsteck, Johannes; Witter, Julia** (2011): Generation of time-consistent industry codes in the face of classification changes: Simple heuristic based on the Establishment History Panel (BHP). FDZ-Methodenreport, 05/2011 (en), Nürnberg.

**Fackler, Daniel; Hank, Eva; Müller, Steffen; Stegmaier, Jens** (2017): Identifying bankruptcies in German social security data. FDZ-Methodenreport, 10/2018 (en), Nürnberg.

**Fitzenberger, Bernd; Osikominu, Aderonke; Völter, Robert** (2006): Imputation rules to improve the education variable in the IAB employment subsample. In: Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, Vol. 126 (3), pp. 405–436.

**Fitzenberger, Bernd; Osikominu, Aderonke; Völter, Robert** (2005): Imputation rules to improve the education variable in the IAB employment subsample. FDZ-Methodenreport, 03/2005 (en), Nürnberg.

**Ganzer, Andreas; Schmucker, Alexandra; Stegmaier, Jens; Stüber, Heiko** (2022): Establishment-History-Panel 1975-2020. FDZ-Datenreport 03/2022 (en), Nürnberg. DOI:10.5164/IAB.FDZD.2203.en.v1

**Hethey, Tanja; Schmieder, Johannes F.** (2010): Using worker flows in the analysis of establishment turnover - Evidence from German administrative data. FDZ-Methodenreport, 06/2010 (en), Nürnberg.

**Hethey-Maier, Tanja; Schmieder, Johannes F.** (2013): Does the use of worker flows improve the analysis of establishment turnover? Evidence from German administrative data. In: Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, Vol. 133 (4), pp. 477–510.

**Lochner, Ben; Seth, Stefan; Wolter, Stefanie** (2023): AKM effects for German labour market data 1985-2021. FDZ-Methodenreport 01/2023 (en), Nuremberg. DOI:10.5164/IAB.FDZM.2301.en.v1

**Ludsteck, Johannes; Thomsen, Ulrich** (2016): Imputation of the Working Time Information for the Employment Register Data. FDZ-Methodenreport 01/2016 (en), Nürnberg.

**Paulus, Wiebke; Matthes, Britta** (2013a): Klassifikation der Berufe. Struktur, Codierung und Umsteigeschlüssel. FDZ-Methodenreport 08/2013 (de), Nürnberg.

**Paulus, Wiebke; Matthes, Britta** (2013b): The German classification of occupations 2010. Structure, coding and conversion table. FDZ-Methodenreport, 08/2013 (en), Nürnberg.

**Schmucker, Alexandra; Ganzer, Andreas; Stegmaier, Jens; Wolter, Stefanie** (2018): Establishment-History-Panel 1975-2017. FDZ-Datenreport 09/2018 (en), Nürnberg. DOI:10.5164/IAB.FDZD.1809.en.v1

**Statistisches Bundesamt (Ed.)** (2003): Klassifikation der Wirtschaftszweige. Edition 2003, Wiesbaden. URL:

https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/klassifikation-wz-2003.html?nn=205976 (Accessed: 30.11.2022, only available in German).

**Statistisches Bundesamt (Ed.)** (2009): Klassifikation der Wirtschaftszweige. 2008, Wiesbaden.URL: https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/klassifikation-wz-2008.html (Accessed: 30.11.2022, only available in German).

**Thomsen, Ulrich; Ludsteck, Johannes; Schmucker, Alexandra** (2018) Skilled or unskilled – Improving the information on qualification for employee data in the IAB Employee Biography. FDZ-Methodenreport, 09/2018 (en), Nürnberg.

**Wermter, Winfried; Cramer, Ulrich** (1988): Wie hoch war der Beschäftigtenanstieg seit 1983? – Ein Diskussionsbeitrag aus der Sicht der Beschäftigtenstatistik der Bundesanstalt für Arbeit. In: Mitteilungen aus der Arbeitsmarkt- und Berufsforschung 4/88, pp. 468–482 (only available in German).

# 7   Table of acronyms

ALWA    Arbeiten und Lernen im Wandel (Working and Learning in a Changing World)

BA    Federal Employment Agency (Bundesagentur für Arbeit, formerly: Bundesanstalt für Arbeit)

BeH    Beschäftigten-Historik des IAB (Employee History File of the IAB)

BHP    Betriebs-Historik-Panel (Establishment History Panel)

DEÜV    Datenerfassungs- und -übermittlungsverordnung (in Kraft getreten am 1. Januar 1999)
(Data Collection and Transmission Regulation (effective as of 1 January 1999)

DEVO    Datenerfassungsverordnung (2. DEVO am 1. Januar 1999 durch die Datenerfassungs- und -
übermittlungsverordnung abgelöst)
(Data Collection Regulation (2nd DEVO was substituted by the Data Collection and
Transmission Regulation on 1 January 1999)

DÜVO    Datenübermittlungsverordnung (2. DÜVO am 1. Januar 1999 durch die Datenerfassungs-
und -übermittlungsverordnung abgelöst)
(Data Transmission Regulation (2nd DÜVO was substituted by the Data Collection and
Transmission Regulation on 1 January 1999)

FDZ    Forschungsdatenzentrum (Research Data Centre)

IAB    Institut für Arbeitsmarkt- und Berufsforschung (Institute for Employment Research)

ISIC    International Standard Industrial Classification of All Economic Activities

KldB    Klassifikation der Berufe (German Classification of Occupations)

NACE    Nomenclature of economic activities (Statistical Classification of Economic Activities in the
European Community)

NUTS    Nomenclature des unités territoriales statistiques (Nomenclature of territorial units for
statistics in the European Community)

SGB    Sozialgesetzbuch (German Social Code)

VZ    Vollzeit (Full-Time)

TS    Tätigkeitsschlüssel (Occupation code)

TZ    Teilzeit (Part-time)

# 8 Appendix

## 8.1 Imputation of information on earnings above the upper earnings limit

Approx. 10% of the information on earnings for people in full-time employment is censored at the upper earnings limit for social security contributions. This leads to bias due to aggregation. Means of wages are biased if the censored observations are not included in the calculation or if censored values are replaced by the censoring limit. No bias occurs for wage quantiles below the censoring limit. As the shares of censored wages can sometimes be very large (well over 50%) depending on the company wage level, in many analyses it would only be possible to use quantiles below the median. In order to remedy these two problems, the information on earnings (average daily wage) was imputed before the statistics (means and medians) were calculated. Our implementation follows that of Card, Heining and Kline (2015).

In order to depict the heterogeneity of wages over time, establishments and qualification levels as extensively as possible, Tobit models (dependent variable: logarithmised daily wage) are estimated separately for cells resulting from the interaction of year (1975–2021), qualification level (3 groups), age (6 groups) and gender (2 groups), whereby the three upper age groups are combined in the group of highly qualified workers[7]. This makes a total of 42 x 2 x (2 x 6 + 1 x 4) = 1, 344 cells or separate Tobit regressions. The explanatory variables are age, indicator for German nationality, east-west indicator, logarithmised establishment size, logarithmised establishment size squared, indicator for one-person establishment, indicator for "large" establishments (10 or more full-time employees), share of censored observations in the establishment excluding the observation currently being imputed ("leave one out"), share of highly qualified employees in the establishment, share of unskilled employees in the establishment. The heterogeneity between establishments and fixed establishment effects are depicted by including the establishment-specific leave-one-out means of the log wages as an additional regressor. (The leave-one-out mean contains only the wages of the colleagues in the respective establishment and ignores the wage of the employee whose wage is being imputed.) The leave-one-out mean is also biased if it contains censored observations. In order to reduce this bias, the imputation is performed in two steps. The first imputation is based on a Tobit model that does not include the leave-one-out means. The imputations from this step are only used to calculate the leave-one-out means. The second step then includes the leave-one-out means as an additional regressor and delivers the final imputations.

The imputed wage for observation $i$ is formed (at each step) by adding a normally distributed residual term $\varepsilon$ to the linear prediction $x_i\,\hat{b}$ from the Tobit model. The standard deviation $\sigma$ of the residual term is equivalent to the estimated variance of the residual of the Tobit model. In order to guarantee that all the imputed values are above the upper earnings limit (valid in the respective year), the residual term is drawn from a correspondingly truncated normal distribution. If $c$ is the censoring limit, $k = \Phi(\frac{c-x_i b}{\sigma})$, where $\Phi(\cdot)$ indicates the distribution function of the standard

---

[7] The following three educational groups (based on educational imputation according to Fitzenberger et al. 2005) are grouped together: 0.1.3; 2.4 and 5.6. The age groups 15-24, 25-34, 35-44, 45-54, 55-64 and 65 and older are formed.

normal distribution, and $u \sim U[0; 1]$ is a uniformly distributed random variable in the interval $[0; 1]$, then the residual $\varepsilon$ is formed in accordance with $\varepsilon = \sigma\, \Phi^{-1}(k + (1 - k)\, u)$.

## 8.2 Person groups in the BeH

When the notification procedure was changed in 1999, the person group code was introduced. It indicates features of the employment relationship that are relevant with regard to social security and benefits legislation and involve at least one branch of social security. Employees subject to social security whose employment relationships have no such special features are given the code 101. If the employment relationship has special features, the codes 102 ff apply. For notifications before 1999 the variable was generated artificially. In this process, notifications with an unclear classification were always given the code "101". Additional person group codes were included later on (see Table 8).

**Table 8: Person group codes in the BeH**

| Code | Name | reported from |
|---|---|---|
| 101 | Employees subject to social security with no special features | 01.01.1999 |
| 102 | Trainees / apprentices with no special features | 01.01.1999 |
| 103 | Employees in partial retirement | 01.01.1999 |
| 104 | Freelance home workers | 01.01.1999 |
| 105 | Interns | 01.01.1999 |
| 106 | Student trainees | 01.01.1999 |
| 108 | Recipients of early retirement benefit | 01.01.1999 |
| 109 | Marginal part-time employees in accordance with § 8 para. 1 No. 1 Social Code Book IV (SGB IV) | 01.01.1999 |
| 110 | Short-term employees in accordance with § 8 para. 1 No. 2 SGB IV | 01.01.1999 |
| 112 | Family workers in agriculture | 01.01.1999 |
| 113 | Part-time farmers | 01.01.2011 |
| 114 | Part-time farmers - employed seasonally | 01.01.2011 |
| 116 | Recipients of compensation according to the Act on Support in the Case of Termination of Farming Activities (FELEG) | 01.01.1999 |
| 118 | Casual workers | 01.01.1999 |
| 119 | Old-age pensioners exempt from insurance contributions and recipients of old-age pension benefits | 01.01.1999 |
| 120 | Persons who are presumed to be in employment (§7 para. 4 SGB IV) | 01.01.1999 |
| 121 | Trainees / apprentices (earnings not above the low-wage threshold) | 01.01.2011 |
| 122 | Trainees / apprentices (external institution) | 01.01.2011 |
| 123 | Persons completing a year of voluntary social or environmental work or Federal Voluntary Service | 01.01.2011 |
| 124 | Home workers | 01.01.1999 |
| 140 | Seamen | 01.01.1999 |
| 141 | Trainees/apprentices in seafaring occupations with no special features | 01.01.1999 |
| 142 | Seamen in partial retirement | 01.01.1999 |
| 143 | Maritime pilots | 01.01.1999 |
| 144 | Trainees/apprentices in seafaring occupations (earnings not above the low-wage threshold) | 01.01.2011 |

| Code | Name | reported from |
|------|------|---------------|
| **149** | Old-age pensioners exempt from insurance contributions and recipients of old-age pension benefits employed in seafaring occupations | 01.01.2005 |
| **150** | Retired persons with compulsory social security payments (mar. shipping) | 01.01.2017 |
| **190** | Employees who are insured solely in the statutory accident insurance | 01.06.2009 |
| **201** | Employees in private households (reported via the "household cheque procedure") | 01.01.1999 |
| **202** | Short-term employees | 01.01.1999 |
| **203** | Artists and publicists subject to social security | 01.01.1999 |
| **205** | Casual workers | 01.01.1999 |
| **207** | Nurses in the sense of § 19 SGB XI/with no eligibility for financial assistance on the part of the person receiving nursing care | 01.01.1999 |
| **208** | Nurses in the sense of § 19 SGB XI/with eligibility for financial assistance on the part of the person receiving nursing care | 01.01.1999 |
| **209** | Marginal part-time employees in private households (reported via the "household cheque procedure") | 01.01.1999 |
| **210** | Short-term employees in private households (reported via the "household cheque procedure") | 01.01.1999 |
| **301** | Persons performing basic military service or voluntary military service | 01.01.1999 |
| **302** | Persons performing reserve duty | 01.01.1999 |
| **303** | Persons performing alternative civilian service | 01.01.1999 |
| **304** | Persons completing a year of voluntary social or environmental work instead of alternative civilian service | 01.06.2003 |
| **305** | Persons performing military service, special types | 01.06.2009 |
| **306** | Special assignment abroad | 01.01.1999 |
| **599** | Other workers | 01.01.1999 |
| **999** | No details available | 01.01.1999 |
| **XXX** | No allocation possible | 01.01.1999 |
| **YYY** | Error in original value | 01.01.1999 |
| **ZZZ** | No details reported | 01.01.1999 |

## 8.3   Frequency tables and labels

Frequency tables and overviews of the individual values and labels of the variables can be found in separate files under:

- https://doku.iab.de/fdz/reporte/2022/DR_12-22_frequencies_labels_de.zip
- https://doku.iab.de/fdz/reporte/2022/DR_12-22_frequencies_labels_en.zip

# Imprint

**Corresponding author**
Alexandra Schmucker
Phone: +49 911 179-1752
Email: alexandra.schmucker@iab.de