Discussion Papers in Economics

No. 01/2017

Administrative Wage and Labor Market Flow Panel (*AWFP*) 1975–2014

Stefan Seth
Institute for Employment Research (IAB)

Heiko Stüber
University of Erlangen-Nürnberg and IAB

ISSN 1867-6707

Friedrich-Alexander-Universität Erlangen-Nürnberg
Institute for Economics
https://www.iwf.rw.fau.de/research/iwf-discussion-paper-series/

Administrative Wage and Labor Market Flow Panel (*AWFP*) 1975–2014

Stefan Seth, Institute for Employment Research (IAB)
Heiko Stüber, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) and IAB

Abstract

This paper describes the new *Administrative Wage and Labor Market Flow Panel* (*AWFP*). The *AWFP* is a dataset on labor market flows and stocks for the universe of German establishments. It contains data on job flows, worker flows, and about wages for each establishment. The *AWFP* contains this information also for partitions of the labor force according to various employee characteristics and for some subgroups of employees. The *AWFP* covers the time period 1975–2014. All data are available at an annual and quarterly frequency.

Keywords: establishment data, job flows, worker flows, wages, German administrative data

Dataset: Version 1.0 (AWFP 1975-2014, v1.0)

Documentation: Version 1.0

Preparation of the dataset: Stefan Seth (Institute for Employment Research, IAB)

Conceptual design: C. Bayer (Universität Bonn), C. Merkl (Friedrich-Alexander-Universität Erlangen-Nürnberg, FAU), S. Seth (IAB), H. Stüber (FAU & IAB), and F. M. Wellschmied (Universidad Carlos III de Madrid)

Acknowledgements

The basic data generation process of the *AWFP* is identical to the one of the recent Establishment History Panel (BHP). Therefore, some sections of this data report (all marked with an asterisk, *) are copied (and only slightly altered) from the data report of the Establishment History Panel (Schmucker et al. 2016). We would like to thank Alexander Schmucker, Johannes Ludsteck, Johanna Eberle, and Andreas Ganzer for the permission to do so.

The authors would like to thank the German Science Foundation (DFG) for financial support.

Contents

Note: Some sections of this data report (all marked with an asterisk, *) are copied (and only slightly altered) from the data report of the Establishment History Panel (Schmucker et al. 2016).

1 Introduction and outline

1.1 Introduction

The *Administrative Wage and Labor Market Flow Panel* (*AWFP*) was generated within the framework of the "Custom Shaped Administrative Data for the Analysis of Labour Market" (CADAL) project and the "Wages, Heterogeneities, and Labor Market Dynamics" project. Both projects are part of the priority program "The German Labor Market in a Globalized World" (SPP 1764), which is sponsored by the German Science Foundation (DFG).

The *AWFP* is a dataset on labor market flows and stocks for the universe of German establishments. It contains data on job flows (changes in the number of employees per establishment), worker flows (information about the hiring and firing activity), and wages for each establishment. The *AWFP* contains this information also for partitions of the labor force according to various employee characteristics (such as sex, education, age, and tenure) and for some sub-groups of employees (e.g., newly hired workers). The *AWFP* covers the time period 1975–2014. All data are available at an annual and quarterly frequency.[1]

The main data source of the *AWFP* data is the Employment History (*Beschäftigten-Historik*, BeH) of the Institute for Employment Research (IAB). The BeH comprises all individuals who were at least once employed subject to social security in Germany since 1975.[2] Some data packages – concerning flows from or into unemployment – use additional data from the Benefit Recipient History (*Leistungsempfängerhistorik*, LeH). The LeH comprises, inter alia, all individuals who received benefits in accordance with Social Code Book III (recorded from 1975 onwards).

With the next update (scheduled for the end of 2017), the *AWFP* data will encompass aggregated *public release* data in addition to the establishment level panel data. This public release data will contain, e.g., information of job and worker flows for groups of firms and thus can be used to study the cyclical dynamics (of fractions) of the labor market in terms of turnover and churning.

1.2 Data access

Certain packages of the *AWFP* will be available as an extension for the next Establishment History Panel (BHP) of the Research Data Centre (FDZ) of the German Federal Employment Agency at the IAB, scheduled to be available in 2018.[3]

The availability of packages will be announced with the next update of the *AWFP*, scheduled for the end of 2017. At the same time we will publish a FDZ data report introducing the *AWFP* extension for the BHP. The aggregated public release data will also be available with the next update of the *AWFP*. It will be downloadable from the IAB website (http://www.iab.de/en) and the website of the Chair of Macroeconomics at the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) (http://www.makro.wiso.fau.de/).

[1] Some data packages will be available on a monthly frequency with the next update of the *AWFP*.
[2] The BeH also comprises marginal part-time workers employed since 1999.
[3] For a data report on the recent BHP see Schmucker et al. 2016.

The BHP data and the *AWFP* extension may be analyzed in the context of a research visit at the FDZ and subsequent 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.

1.3 Outline

Categories	Descriptions
Topics	• General establishment data (e.g., location and industry, wages). • Stratified employment stock data (e.g., qualification of workers). • Stratified inflow and outflow data (e.g., from and to unemployment).
Research unit	Establishments in Germany with at least one full-time employee subject to social security.
Number of cases	Annual number of observations: 1.4–3.3 million establishments
Period covered	West Germany: 1975–2014 East Germany: 1992–2014
Time reference	Annual frequency: 31 December of each year Quarterly frequency: last day of each quarter Monthly frequency: last day of each month
Regional structure	Districts (*Kreise*)
Type of territorial allocation	Corrected territorial allocation as of 31/12/2014
Frequency of data collection	Annual frequency Quarterly frequency
File format	Stata
Data access	On-site use or remote data access

2 Data sources

2.1 Employee history (BeH)*

The source of data regarding employment is the Employee History (*Beschäftigten-Historik*, BeH, V10.0.0) 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 (for further details see Bender et al. 1996 and Wermter/Cramer 1988). Under this so called DEÜV procedure (previously DEVO/DÜVO) 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 observation period of the BeH V10.0.0 extends from 1 January 1975 to 31 December 2014.

2.1.1 The establishment concept*

The concept of an establishment is central to the *AWFP*. The information about establishments contained in the BeH 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).

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 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.

Once an establishment has been allocated an establishment number it remains unchanged in principle (see Schmucker et al. 2016 for more detailed information and exceptions).

2.2 Benefit recipient history (LeH)

Some data packages – concerning flows from or into unemployment – use additional data from the Benefit Recipient History (*Leistungsempfängerhistorik*, LeH, V7.3.0). The LeH comprises, inter alia, all individuals that received benefits in accordance with Social Code Book III (recorded from 1975 onwards).

3 Data preparation – corrections and validation procedures performed on the micro-level data

For compiling the *AWFP*, the employment notifications of the BeH (see section 2.1) are aggregated at establishment level using the establishment ID (see section 2.1.1). Before the aggregation, the data on individuals are subjected to numerous validation procedures.

3.1 Selection of the notifications in the BeH*

The data on individuals from the BeH are used as the basis for the *AWFP*, 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, 118, 119, 120, 121, 122, 140, 141, 142, 143, 144, 149, 201, 203, 205, 209, 999, YYY (see Appendix).
- Notifications with a wage of 0 are deleted. As these notifications concern de-registrations 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).

3.2 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 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, recognized 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. 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 and 2006) and enhanced by Kruppe et al. (2014). 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. For more information on the imputation please refer to section 8.1 of Schmucker et al. (2016).

3.3 Validation of the information on earnings

3.3.1 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 de-registrations. 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 *AWFP*.

3.3.2 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.

3.3.3 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. This leads to biased results due to aggregation because 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 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 8.2 of Schmucker et al. (2016).

3.4 Validation of the information on full-time and part-time employment*

For a transitional period after the introduction of the new occupation code in December 2011 it was permitted to leave out the information on the occupation code and working time in the social security notifications. In a good 10% of the notifications submitted by the establishments between December 1st 2011 and May 31th 2012 the information regarding working hours is therefore missing. For this reason a logit model was developed at the IAB which can be used to impute the missing information (see Ludsteck and Thomsen 2016). The information on working hours that is generated using this procedure is contained in this variable.

3.5 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.

4 Data quality

The data quality of the *AWFP* depends on the data quality of the underlying BeH data, which we discuss below.

4.1 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.

4.2 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. This is done using files of employment notifications for one particular year which were submitted 36, 30, 18, 12 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 2012 it was possible to use a 30-month file, for 2013 an 18-month file and for 2014 (only) a 6-month file. It can be assumed that the number of establishments is slightly under-recorded for the years 2012 and 2013. It can also be assumed that there are larger gaps for 2014, which makes it advisable to compare the 6-, 12- and 18-month files for 2013: for instance the 12-month file contains 0.8% more employees than the 6-month file. At establishment level the notifications that were submitted late had a stronger effect: after 12 months an additional 2.6% of the establishments are recorded. What is noticeable here is that most of these establishments are very small establishments with up to ten employees. Although the number of employees increases again by 1.3% between the 12-month and the 18-month files, the increase recorded in the number of establishments is only 0.5%. During this period more establishments with more than 200 employees were added to the data.

4.3 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).

4.4 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 filed more correctly.

4.5 Classification of economic activities

During the observation period of the *AWFP* the classification of economic activities has changed several times. This makes longitudinal analysis difficult. The FDZ developed a method to impute time-consistent industry codes (see Eberle et al. 2011). The *AWFP* therefore includes four original establishment industry classifications (w73, w93, w03, and w08) and two imputed classifications (w73_imp and w93_imp).

More information on the classifications is provided by the German Federal Statistical Office (https://www.destatis.de/DE/Methoden/Klassifikationen/Klassifikationen.html) and the German Federal Employment Agency (Bundesagentur für Arbeit 2010, Bundesanstalt für Arbeit 1973 and 1996).

5 Generating the Administrative Wage and Labor Market Flow Panel (*AWFP*)

5.1 Overview of the dataset

The *AWFP* data are divided into data packages. These packages can be grouped into three categories:

1) packages containing general establishment data,
2) packages containing (stratified) employment stock data, and
3) packages containing stratified employment flow data.

All packages, p, of the *AWFP* are saved using the following structure: p*Tn*, where *T* indicates the frequency of the data (a = annual, q = quarter, m = month) and *n* indicates the package number.[4]

[4] Packages at the monthly frequency will be available with the next update of the *AWFP*.

In order to minimize the memory requirements of the *AWFP*:

- Some variables are not explicitly included in the data if they can be calculated using the available information (see section 5.3).
- Observations are not included in packages if all variables of the package (excluding establishment and time identifier) were zero or missing. Therefore certain missing values that will be generated while merging packages should be replaced by zeros (see section 5.4 and programming example in section 5.5).
- Most stock packages are only available on the quarterly frequency, since the stock for the 4th quarter of a year corresponds to the stock at the yearly frequency (see section 5.5 for a programming example).

After the data on individuals have been preprocessed (see section 3) the packages are generated as follows:

- Selection of all BeH observations that include the respective reference date.
- 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 job is recorded for one person in the same establishment, the job with the higher daily wage is selected.
- Aggregation of all employment notifications as of the reference date to form selected statistics at establishment level on the basis of the establishment ID.

The stocks and flows in the *AWFP* are generally calculated on a 'regular worker' basis. In the next section we will define the notion 'regular worker' and give our standard definition of how we calculate stocks and flows. Unless explicitly mentioned otherwise these standard definitions are used for the generation of the *AWFP*.

5.2 Definitions

All data packages – except package p_101 – contain information at an annual (a) or quarterly (q) frequency. Hence, when we talk about a "period", we think of a year or a quarter.

5.2.1 Regular workers

We define a person as a 'regular worker' when he/she is full-time employed and belongs to person group 101 (employees s.t. social security without special features), 140 (seamen) or 143 (maritime pilots) in the BeH. Therefore all (marginal) part-time employees, employees in partial retirement, interns etc. are not accounted for as regular workers. See the appendix for more details on the person group in the BeH. The stocks and flows in the *AWFP* are generally calculated on a 'regular worker' basis.

5.2.2 Normal workers

Some packages contain information on 'normal' workers. Normal workers are defined like regular workers (see above) but they may work part-time. Therefore each regular worker is also a normal worker but not vice versa.

5.2.3 Other workers

Some packages contain information on 'other' workers. Other workers are neither normal workers, apprentices, workers in partial retirement nor workers in marginal part-time. This group consists mainly of interns *(Praktikanten/Werkstudenten)*.

5.2.4 Stocks

The **stock** of employees of an establishment in some period *t* equals the number of employees on the last day of period *t*. Unless explicitly mentioned otherwise, we calculate stocks based on regular workers and using the 'end-of-period' definition.

Several stocks are broken down according to various characteristics such as age groups. Further information on the individual variables can be found in section 6.

5.2.5 Flows

Inflows of employees of an establishment for period *t* equals the number of employees who were regularly employed on the last day of period *t* but were not on the last day of the preceding period, *t*-1.

Outflows of employees of an establishment for period *t* equals the number of employees who were regularly employed on the last day of the preceding period (*t-1*) but were not on the last day of period *t*.

Unless explicitly mentioned otherwise, we calculate both inflows and outflows based on regular workers and using the 'end-of-period' definition.

Employees who join an establishment and leave it again between two reference dates are not recorded by this flow concept.

Note that a worker counted as an inflow is not necessarily a new hire. For instance, an apprentice who becomes a regular worker represents an inflow because an apprentice is not a regular worker. Analogously, a worker counted as an outflow might remain employed in the same establishment. A regular worker who, for instance, reduces hours and changes to a part-time job represents an outflow. The "status change" package (312, 412) informs about these kinds of flows.

Like some stocks, several inflows and outflows are broken down according to various characteristics such as age groups. Further information on the individual variables can be found in section 6.

5.3 Calculation of omitted variables

In order to minimize the memory requirement of the *AWFP*, some variables that the users can calculate themselves from the available information are not included in the data. For example, the stock of female workers can be calculated as the number of all workers minus the number of male workers (st_female = st_eop - st_male). The tables, in section 6, of these variables are colored grey.

5.4 Merging packages

When merging packages take the following particularities of the *AWFP* data into account:

- Package p_101 contains all establishments in the BeH. Some of them won't show up in other packages, because some establishments existed only for a very short time (between two reference dates).
- Establishments with no inflows or outflows in some period do not appear in the corresponding package for that certain period. However, the establishment does appear for that certain period, e.g., in the package containing wage information (p*105) if the establishment employs at least one regular worker. After merging the packages, the inflow / outflow information for the firm in the certain period will be missing (.) and needs to be replace by zero (0).
- Establishments that have closed down appear only in the worker outflow dataset in the following period. The outflows listed here are equivalent to the employee stocks of the preceding period (period of exit).

Inflows / outflows cannot be calculated for the first / last available period due to missing values of the preceding / next period.

5.5 Programming example

Example 1: The following Stata programming example shows how a flow panel dataset on the quarterly frequency can be created. Missing values (.) generated due to the structure of the *AWFP* (see section 5.4) are replaced by zero (0).

```
use "$orig/pq207", clear // package containing the stock of regular workers
merge 1:1 betnr q using "$orig/pq307", nogen // package containing inflows
merge 1:1 betnr q using "$orig/pq407", nogen // package containing outflows
gen byte a = ceil(q/4) // generating year index
* Merging general (time inconsistent) establishment data
merge m:1 betnr using "$orig/p_101" // package containing dates (e.g. first appearance of establishment)
keep if _merge == 3 // drop establishments that are contained in p_101 only
drop _merge
merge m:1 betnr a using "$orig/pa102", nogen keep(3) // package containing location and industry
```

* Replace missing values (.) that originate from merging with 0:
replace in_eop = 0 if in_eop == . // inflows
replace in_male = 0 if in_male == . // male inflows
replace out_eop = 0 if out_eop == . // outflows
replace out_male = 0 if out_male == . // male outflows

Example 2: The following Stata programming example shows how create a dataset on the annual frequency using a quarterly frequency stock dataset (see also section 5.1).

```
use "$orig/pq207" if mod(q,4) == 0, clear // package containing the
stock of regular workers, only the 4th quarters are used
gen byte a = ceil(q/4) // generating year index
label variable a "year index"
drop q
save "$data/pa207", replace
```

6 Description of the variables and characteristics

Remember, unless explicitly mentioned otherwise, we calculate all stock, inflows, and outflows based on regular workers and using the 'end-of-period' definition (see section 5.2)!

6.1 Common identifiers

All datasets contain the establishment identifier and – except package p_101 – one time index. The second letter of the filename indicates which time index is included in the package (a = year, q = quarter, m = month).

6.1.1 Establishment identifier (betnr)

Variable name	Betnr
Origin	Generated variable
Detailed description	Identifies the observation unit (plant/establishment) across time and packages.

6.1.2 Index of year (a)

Variable name	A
Origin	
Detailed description	Starts with 1, with year no 1 being the year 1975.

6.1.3 Index of quarter (q)

Variable name	q
Origin	
Detailed description	Starts with 1, with quarter no 1 being the first quarter of 1975.

6.2 Package p_101: dates

Note that 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.).

6.2.1 Foundation date (founded_m)

Variable name	founded_m
Origin	
Detailed description	The first month in which the establishment has an employee; possibly left-censored.

6.2.2 First regular worker (first_rw_m)

Variable name	first_rw_m
Origin	
Detailed description	The first month in which the establishment has employed a 'regular worker' (as defined above); the value is missing if there has not been such a month; possibly left-censored.

6.2.3 Shut down date (shut_m)

Variable name	shut_m
Origin	
Detailed description	The last month at which the establishment had an employee; possibly right-censored.

6.2.4 Last regular worker (last_rw_m)

Variable name	last_rw_m
Origin	
Detailed description	The last month in which the establishment employed a 'regular worker' (as defined in section 5.2.1); missing values occur; possible right-censored.

Note: To transform the month information of the variables in package p_101 into quarter or year information, the following codes can be used:

Quarter: gen `var'_q = ceil(`var'_m/3)

Year: gen `var'_a = ceil(`var'_m/12)

6.3 Package pa102: location and industry

6.3.1 Establishment location (district)

Variable name	district
Origin	
Detailed description	The district (Kreis) the establishment is located.

6.3.2 Establishment's industry classification (w73)

Variable name	w73
Origin	
Detailed description	Establishment's industry classification according to the German Classification of Economic Activities WZ 73; filled 1975–2002; missing values occur.

6.3.3 Establishment's industry classification (w93)

Variable name	w93
Origin Detailed description	Establishment's industry classification according to the German Classification of Economic Activities WZ 93; filled 1998–2003; missing values occur.

6.3.4 Establishment's industry classification (w03)

Variable name	w03
Origin Detailed description	Establishment's industry classification according to the German Classification of Economic Activities WZ 03; filled 2003–2008; missing values occur.

6.3.5 Establishment's industry classification (w08)

Variable name	w08
Origin Detailed description	Establishment's industry classification according to the German Classification of Economic Activities WZ 08; filled since 2008; missing values occur.

6.3.6 Establishment's industry classification (w73_imp)

Variable name	w73_imp
Origin Detailed description	Imputed / transcoded establishment's industry classification according to the German Classification of Economic Activities WZ 73

6.3.7 Establishment's industry classification (w93_imp)

Variable name	w93_imp
Origin Detailed description	Imputed / transcoded establishment's industry classification according to the German Classification of Economic Activities WZ 93

6.4 Package 103: age

6.4.1 Mean age (mean_age)

Variable name	mean_age
Origin Detailed description	The mean age within the establishment of workers at the end of the period (in years).

6.5 Package 104: tenure

6.5.1 Mean tenure (mean_tenure)

Variable name	mean_tenure
Origin Detailed description	The mean tenure within the establishment of workers at the end of the period (in quarters); possibly left-censored.

6.6 Package 105: wages (regular workers, stayers, inflows, outflows)

An addition to the package name indicates the underlying universe of the package: _all = all regular workers; _in = new regular workers (inflows); _out = outgoing regular workers (outflows); _st = incumbent regular workers (stayers).

Notes: All wage information might be right-censored.

6.6.1 Number of observations (num_all, num_st, num_in, num_out)

Variable name	num_all, num_st, num_in, num_out
Origin Detailed description	Number of workers the calculation is based on.

6.6.2 Mean wage of workers (dw_mean)

Variable name	dw_mean
Origin Detailed description	Mean daily wage of all/incumbent/new/outgoing workers at the end of the period. Wages of outflows are calculated with respect to the preceding period.

6.6.3 Standard deviation of wage (dw_sd)

Variable name	dw_sd
Origin Detailed description	Standard deviation of daily wages of all/incumbent/new/outgoing workers at the end of the period. Wages of outflows are calculated with respect to the preceding period.

6.6.4 25th percentile of wage (dw_p25)

Variable name	dw_p25
Origin Detailed description	25th percentile of the daily wage of all/incumbent/new/outgoing workers at the end of the period. Wages of outflows are calculated with respect to the preceding period.

6.6.5 50th percentile of wage (dw_p50)

Variable name	dw_p50
Origin Detailed description	Median daily wage of all/incumbent/new/outgoing workers at the end of the period. Wages of outflows are calculated with respect to the preceding period.

6.6.6 75th percentile of wage (dw_p75)

Variable name	dw_p75
Origin Detailed description	75th percentile of the daily wage of all/incumbent/new/outgoing workers at the end of the period. Wages of outflows are calculated with respect to the preceding period.

6.7 Package 105imp: Wage

Notes: It is the same as Package 105 (see section 6.6), but using imputed wages (see section 3.3.3). Hence, not right-censored.

6.8 Package 203: Age

6.8.1 Stock of workers aged 15–17 (st_age_1)

Variable name	st_age_1
Origin Detailed description	Stock of workers with a certain age at the end of the period.

6.8.2 Stock of workers aged 18–24 (st_age_2)

Variable name	st_age_2
Origin Detailed description	Stock of workers with a certain age at the end of the period.

6.8.3 Stock of workers aged 25–29 (st_age_3)

Variable name	st_age_3
Origin Detailed description	Stock of workers with a certain age at the end of the period.

6.8.4 Stock of workers aged 30–44 (st_age_4)

Variable name	st_age_4
Origin Detailed description	Stock of workers with a certain age at the end of the period.

6.8.5 Stock of workers aged 45–49 (st_age_5)

Variable name	st_age_5
Origin Detailed description	Stock of workers with a certain age at the end of the period.

6.8.6 Stock of workers aged 50–54 (st_age_6)

Variable name	st_age_6
Origin Detailed description	Stock of workers with a certain age at the end of the period.

6.8.7 Stock of workers aged 55–59 (st_age_7)

Variable name	st_age_7
Origin Detailed description	Stock of workers with a certain age at the end of the period.

6.8.8 Stock of workers aged 60 and older (st_age_8)

Variable name	st_age_8
Origin Detailed description	Stock of workers with a certain age at the end of the period.

6.9 Package 204: tenure

6.9.1 Stock of workers with a job tenure of up to 1 quarter (st_senio_1)

Variable name	st_senio_1
Origin Detailed description	Number of workers at the end of the period who have been working in the establishment for a certain time.

6.9.2 Stock of workers with a job tenure of 2–4 quarters (st_senio_2)

Variable name	st_senio_2
Origin Detailed description	Number of workers at the end of the month who have been working in the establishment for a certain time.

6.9.3 Stock of workers with a job tenure of 5 quarters – 3 years (st_senio_3)

Variable name	st_senio_3
Origin Detailed description	Number of workers at the end of the month who have been working in the establishment for a certain time.

6.9.4 Stock of workers with a job tenure of 13 quarters – 9 years (st_senio_4)

Variable name	st_senio_4
Origin Detailed description	Number of workers at the end of the month who have been working in the establishment for a certain time.

6.9.5 Stock of workers with a job tenure of more than 9 years (st_senio_5)

Variable name	st_senio_5
Origin Detailed description	Number of workers at the end of the month who have been working in the establishment for a certain time.

6.10 Package 206: workers – various stock definitions

6.10.1 Stock of workers at the beginning of the period (st_bop)

Variable name	st_bop
Origin Detailed description	Number of workers as of the first day of the period.

6.10.2 Stock of workers according to LEHD flow definition (st_lehd)

Variable name	st_lehd
Origin Detailed description	Number of workers employed for at least one day in the current period.

6.10.3 Stock of workers according to LEHD full-period definition (st_lehd_fp)

Variable name	st_lehd_fp
Origin Detailed description	Number of workers employed for at least one day in the current period, at least one day in the preceding period, and at least one day in the subsequent period.

6.11 Package 207: sex and hours of work

6.11.1 Stock of workers (st_eop)

Variable name	st_eop
Origin Detailed description	Number of workers as of the last day of the period (end-of-period employment).

6.11.2 Stock of male workers (st_male)

Variable name	st_male
Origin Detailed description	Stock of regular male workers at the end of the month.

6.11.3 Stock of female workers

Variable name	
Origin Detailed description	Computable. Stock of regular female workers at the end of the period. Can be calculated as: st_eop – st_male

6.12 Package 208: qualification

6.12.1 Stock of low-skilled workers (st_qual_1)

Variable name	st_qual_1
Origin Detailed description	Stock of workers without formal vocational training (according to the imputed education variable).

6.12.2 Stock of medium-skilled workers (st_qual_2)

Variable name	st_qual_2
Origin Detailed description	Stock of workers with formal vocational training (according to the imputed education variable).

6.12.3 Stock of high-skilled workers (st_qual_3)

Variable name	st_qual_3
Origin Detailed description	Stock of workers with an university degree (according to the imputed education variable).

6.13 Package 211: qualification according to Blossfeld

For information concerning (the validation of) the occupation data, used to form the categories of the Blossfeld classification (see Blossfeld 1987), please refer to sections 4.6 and 3.1.5 of Schmucker et al. (2016).

6.13.1 Stock of workers classified as low-skilled according to Blossfeld (st_task_1)

Variable name	st_task_1
Origin Detailed description	Agricultural occupations, elementary manual occupations, elementary personal services occupations, elementary administrative occupations

6.13.2 Stock of workers classified as medium-skilled according to Blossfeld (st_task_2)

Variable name	st_task_2
Origin Detailed description	Skilled manual occupations, skilled services occupations, skilled administrative occupations

6.13.3 Stock of workers classified as semi-skilled according to Blossfeld (st_task_3)

Variable name	st_task_3
Origin Detailed description	Technicians, associate professionals

6.13.4 Stock of workers classified as high-skilled according to Blossfeld (st_task_4)

Variable name	st_task_4
Origin Detailed description	Professional occupations, managers

6.14 Package 214: apprentices, partial retirement, marginal part-time, and interns (not calculated on a regular worker basis!)

6.14.1 Stock of ‘normal’ workers (st_nml)

Variable name	st_nml
Origin Detailed description	Number of all ‘normal’ workers (see section 5.2.2).

6.14.2 Stock of apprentices (st_app)

Variable name	st_app
Origin Detailed description	Number of apprentices/trainees (*Auszubildende*)

6.14.3 Stock of workers in partial retirement (st_pr)

Variable name	st_pr
Origin Detailed description	Number of workers in partial/progressive retirement (*Altersteilzeit*).

6.14.4 Stock of marginal part-time workers (st_mpt)

Variable name	st_mpt
Origin Detailed description	Number of marginal part-time workers (*geringfügig entlohnte Beschäftigte).*

6.14.5 Stock of 'other' worker (st_other)

Variable name	st_other
Origin Detailed description	Number of 'others' workers (see section 5.2.3).

6.14.6 Stock of all employees

Variable name	
Origin Detailed description	Computable. Number of all employees. Can be calculated as: st_nml + st_app + st_pr + st_mpt + st_other

6.15 Package 303: age

6.15.1 Inflows aged 15–17 years (in_age_1)

Variable name	in_age_1
Origin Detailed description	Inflows by age group. Age is calculated with respect to the end of the period.

6.15.2 Inflows aged 18–24 years (in_age_2)

Variable name	in_age_2
Origin Detailed description	Inflows by age group. Age is calculated with respect to the end of the period.

6.15.3 Inflows aged 25–29 years (in_age_3)

Variable name	in_age_3
Origin Detailed description	Inflows by age group. Age is calculated with respect to the end of the period.

6.15.4 Inflows aged 30–44 years (in_age_4)

Variable name	in_age_4
Origin Detailed description	Inflows by age group. Age is calculated with respect to the end of the period.

6.15.5 Inflows aged 45–49 years (in_age_5)

Variable name	in_age_5
Origin Detailed description	Inflows by age group. Age is calculated with respect to the end of the period.

6.15.6 Inflows aged 50–54 years (in_age_6)

Variable name	in_age_6
Origin Detailed description	Inflows by age group. Age is calculated with respect to the end of the period.

6.15.7 Inflows aged 55–59 years (in_age_7)

Variable name	in_age_7
Origin Detailed description	Inflows by age group. Age is calculated with respect to the end of the period.

6.15.8 Inflows aged 60 years and older (in_age_8)

Variable name	in_age_8
Origin Detailed description	Inflows by age group. Age is calculated with respect to the end of the period.

6.16 Package 306: various inflow definitions

6.16.1 Inflows using daily count (in_dc)

Variable name	in_dc
Origin Detailed description	Number of “entry-events” (a worker employed today but not the preceding day) in the current period.

6.16.2 Inflows according to LEHD definition (in_lehd)

Variable name	in_lehd
Origin Detailed description	Number of workers regularly employed for at least 1 day in the current period but not in the preceding period.

6.16.3 Inflows according to LEHD full-period definition (in_lehd_fp)

Variable name	in_lehd_fp
Origin Detailed description	Number of workers regularly employed for at least 1 day in the current, subsequent and preceding period but not in the period before.

6.17 Package 307: standard-definition inflows and sex

6.17.1 Inflows using the standard end-of-period definition (in_eop)

Variable name	in_eop
Origin Detailed description	Number of regular workers employed at the end of the current period but not employed as regular workers at the end of the preceding period in the same establishment.

6.17.2 Male inflows (in_male)

Variable name	in_male
Origin Detailed description	Number of male inflows.

6.17.3 Female inflows

Variable name	
Origin Detailed description	Computable. Number of female inflows. Can be calculated as: in_eop – in_male

6.18 Package 308: qualification

6.18.1 Inflows of low-skilled workers (in_qual_1)

Variable name	in_qual_1
Origin Detailed description	Number of inflows without formal vocational training

6.18.2 Inflows of medium-skilled workers (in_qual_2)

Variable name	in_qual_2
Origin Detailed description	Number of inflows with formal vocational training

6.18.3 Inflows of high-skilled workers (in_qual_3)

Variable name	in_qual_3
Origin Detailed description	Number of inflows with an university degree.

6.19 Package 309a: flows from unemployment (ALG / ALG1)

Note that the definition of the duration of unemployment depends on the frequency of the package.

6.19.1 Inflows from unemployment I (in_ue1)

Variable name	in_ue1
Origin Detailed description	Number of inflows previously unemployed for 1 quarter (yearly data: 1 year)

6.19.2 Inflows from unemployment II (in_ue2)

Variable name	in_ue2
Origin Detailed description	Number of inflows previously unemployed for 2 quarters (2 years)

6.19.3 Inflows from unemployment III (in_ue3)

Variable name	in_ue3
Origin Detailed description	Number of inflows previously unemployed for 3–4 quarters (3 years)

6.19.4 Inflows from unemployment IV (in_ue4)

Variable name	in_ue4
Origin Detailed description	Number of inflows previously unemployed for 5–12 quarters (4+ years)

6.19.5 Inflows from unemployment V

Variable name	
Origin Detailed description	Computable. Number of inflows previously unemployed for more than 12 quarters. Can be computed as: in_eop - (in_ue1 + in_ue2 + in_ue3 + in_ue4). Yearly data: in_ue5 ≡ 0.

6.20 Package 309b: flows from non-employment (excl. ALG / ALG1)

Note that the definition of the duration of non-employment depends on the frequency of the package.

6.20.1 Inflows from non-employment I (in_oolf1)

Variable name	in_oolf1
Origin Detailed description	Number of inflows previously economically inactive for 1 quarter (1 year)

6.20.2 Inflows from non-employment II (in_oolf2)

Variable name	in_oolf2
Origin Detailed description	Number of inflows previously economically inactive for 2 quarters (2 years)

6.20.3 Inflows from non-employment III (in_oolf3)

Variable name	in_oolf3
Origin Detailed description	Number of inflows previously economically inactive for 3-4 quarters (3 years)

6.20.4 Inflows from non-employment IV (in_oolf4)

Variable name	in_oolf4
Origin Detailed description	Number of inflows previously economically inactive for 5–12 quarters (4+ years)

6.20.5 Inflows from non-employment V

Variable name	in_oolf5
Origin Detailed description	Computable. Number of inflows previously economically inactive for more than 12 quarters. Can be computed as: in_eop - (in_oolf1 + in_oolf2 + in_oolf3 + in_oolf4). Yearly data: in_oolf5 ≡ 0.

6.21 Package 309c: decomposition of inflows

6.21.1 Inflows from unemployment (in_ue)

Variable name	in_ue
Origin Detailed description	Inflows who were unemployed, i. e., receiving unemployment benefits (*Arbeitslosengeld*) at the end of the preceding period.

6.21.2 Inflows from non-employment (in_oolf)

Variable name	in_oolf
Origin Detailed description	Inflows who were economically inactive, i. e. neither employed nor unemployed, at the end of the preceding period.

6.21.3 Inflows from employment (in_e)

Variable name	in_e
Origin Detailed description	Inflows who were employed at the end of the preceding period (i. e., who changed the employer/establishment).

6.21.4 Start of career inflows

Variable name	
Origin Detailed description	Computable. Number of workers regularly employed for the first time in the current period. Can be calculated as: in_eop - (in_ue + in_oolf + in_e)

6.22 Package 310: wage change

Note: Wages might be right-censored. Employees with censored wages in both observation periods will be classified as workers with rigid wages.

6.22.1 Inflows with wage decrease by at least 2 percent (in_dw_dec2)

Variable name	in_dw_dec2
Origin Detailed description	Number of inflows (from employment) experiencing a nominal wage decrease by at least 2 percent.

6.22.2 Inflows with wage decrease by at least 4 percent (in_dw_dec4)

Variable name	in_dw_dec4
Origin Detailed description	Number of inflows (from employment) experiencing a nominal wage decrease by at least 4 percent.

6.22.3 Inflows with wage increase by at least 2 percent (in_dw_inc2)

Variable name	in_dw_inc2
Origin Detailed description	Number of inflows (from employment) experiencing a nominal wage increase by at least 2 percent.

6.22.4 Inflows with wage increase by at least 4 percent (in_dw_inc4)

Variable name	in_dw_inc4
Origin Detailed description	Number of inflows (from employment) experiencing a nominal wage increase by at least 4 percent.

6.22.5 Inflows with an absolute wage change less than 2 percent (in_dw_rig2)

Variable name	in_dw_rig2
Origin Detailed description	Number of inflows (from employment) experiencing nominal wage rigidity, i. e., an absolute wage change less than 2 percent.

6.22.6 Inflows with an absolute wage change less than 4 percent (in_dw_rig4)

Variable name	in_dw_rig4
Origin Detailed description	Number of inflows (from employment) experiencing nominal wage rigidity, i. e., an absolute wage change less than 4 percent.

6.22.7 Inflows from young establishments with wage wage decrease by at least 4 percent (in_dw_dec4_yng)

Variable name	in_dw_dec4_yng
Origin Detailed description	Number of inflows from young establishments experiencing a nominal wage decrease by at least 4 percent. Establishments are classified as young when they have been operating for not more than 3 quarters (quarterly frequency) or 1 year (yearly frequency).

6.22.8 Inflows from young establishments with wage increase by at least 4 percent (in_dw_inc4_yng)

Variable name	in_dw_inc4_yng
Origin Detailed description	Number of inflows from young establishments experiencing a nominal wage increase by at least 4 percent. Establishments are classified as young when they have been operating for not more than 3 quarters (quarterly frequency) or 1 year (yearly frequency).

6.22.9 Inflows from young establishments with an absolute wage change less than 4 percent (in_dw_rig4_yng)

Variable name	in_dw_rig4_yng
Origin Detailed description	Number of inflows from young establishments experiencing nominal wage rigidity, i. e., an absolute wage change less than 4 percent. Establishments are classified as young when they have been operating for not more than 3 quarters (quarterly frequency) or 1 year (yearly frequency).

6.23 Package 311: qualification according to Blossfeld

For information concerning (the validation of) the occupation data, used to form the categories of the Blossfeld classification (see Blossfeld 1987), please refer to sections 4.6 and 3.1.5 of Schmucker et al. (2016).

6.23.1 Inflows classified as low-skilled according to Blossfeld (in_task_1)

Variable name	in_task_1
Origin Detailed description	Agricultural occupations, elementary manual occupations, elementary personal services occupations, elementary administrative occupations

6.23.2 Inflows classified as medium-skilled according to Blossfeld (in_task_2)

Variable name	in_task_2
Origin Detailed description	Skilled manual occupations, skilled services occupations, skilled administrative occupations

6.23.3 Inflows classified as semi-skilled according to Blossfeld (in_task_3)

Variable name	in_task_3
Origin Detailed description	 Technicians, associate professionals

6.23.4 Inflows classified as high-skilled according to Blossfeld (in_task_4)

Variable name	in_task_4
Origin Detailed description	 Professional occupations, managers

6.24 Package 312: status change inflows

6.24.1 External inflows

Variable name	
Origin Detailed description	Computable. Number of inflows (regular workers) who were employed by another establishment at the preceding reference date, but not as regular workers: in_ext = in_app_ext + in_mpt_ext + in_pr_ext + in_other_ext + in_pt_ext

6.24.2 Internal inflows

Variable name	
Origin Detailed description	Computable. Number of regular workers who were employed by the same establishment at the preceding reference date, but not as regular workers. in_int = in_app_int + in_mpt_int + in_pr_int + in_other_int + in_pt_int

6.24.3 External inflows of former apprentices (in_app_ext)

Variable name	in_app_ext
Origin Detailed description	 Number of inflows employed as apprentices in **another establishment** at the preceding reference date.

6.24.4 External inflows of former marginal part-time workers (in_mpt_ext)

Variable name	in_mpt_ext
Origin Detailed description	 Number of inflows employed as marginal part-time workers in **another establishment** at the preceding reference date.

6.24.5 External inflows of workers previously in partial retirement (in_pr_ext)

Variable name	in_pr_ext
Origin Detailed description	 Number of inflows employed as workers in partial retirement in **another establishment** at the preceding reference date.

6.24.6 External inflows of former ‘other’ workers (in_other_ext)

Variable name	in_other_ext
Origin Detailed description	Number of inflows employed as ‘other’ workers (see section 5.2.3) in **another establishment** at the preceding reference date.

6.24.7 External inflows of former 'normal' workers in part-time (in_pt_ext)

Variable name	in_pt_ext
Origin Detailed description	Number of inflows employed as 'normal' workers (see section 5.2.2) in part-time in **another establishment** at the preceding reference date.

6.24.8 Internal inflows of former apprentices (in_app_int)

Variable name	in_app_int
Origin Detailed description	Number of inflows employed as apprentices by **the same establishment** at the preceding reference date.

6.24.9 Internal inflows of former marginal part-time workers (in_mpt_int)

Variable name	in_mpt_int
Origin Detailed description	Number of inflows employed as marginal part-time workers by **the same establishment** at the preceding reference date.

6.24.10 Internal inflows of workers previously in partial retirement (in_pr_int)

Variable name	in_pr_int
Origin Detailed description	Number of inflows employed as workers in partial retirement by **the same establishment** at the preceding reference date.

6.24.11 Internal inflows of former ‘other’ workers (in_other_int)

Variable name	in_other_int
Origin Detailed description	Number of inflows employed as ‘other’ workers (see section 5.2.3) by **the same establishment** at the preceding reference date.

6.24.12 Internal inflows of former 'normal' workers in part-time (in_pt_int)

Variable name	in_pt_int
Origin Detailed description	Number of inflows employed as 'normal' workers (see section 5.2.2) in part-time by **the same establishment** at the preceding reference date.

6.25 Package 313: hires and re-hires

6.25.1 New hires (in_new)

Variable name	in_new
Origin Detailed description	Number of workers not employed (any status!) by the same establishment at the 4 preceding reference dates

6.25.2 Re-hires (in_rehire)

Variable name	in_rehire
Origin Detailed description	Number of workers not employed by same establishment at the preceding reference date but employed (any status!) at one of the 3 reference dates before

6.26 Package 314: apprentices, partial retirement, marginal part-time, and interns (not calculated on a regular worker basis!))

6.26.1 Inflow of normal workers (in_nml)

Variable name	in_nml
Origin Detailed description	Number of 'normal' workers (see section 5.2.2) who were not employed by this establishment in the preceding period.

6.26.2 Inflows of apprentices (in_app)

Variable name	in_app
Origin Detailed description	Number of apprentices/trainees (Auszubildende) who were not employed by this establishment in the preceding period.

6.26.3 Inflow of partial retirement workers (in_pr)

Variable name	in_pr
Origin Detailed description	Number of workers in partial/progressive retirement (Altersteilzeit) who were not employed by this establishment in the preceding period.

6.26.4 Inflow of marginal part-time workers (in_mpt)

Variable name	in_mpt
Origin Detailed description	Number of marginal part-time workers (geringfügig entlohnte Beschäftigte) who were not employed by this establishment in the preceding period.

6.26.5 Inflow of other workers (in_other)

Variable name	in_other
Origin Detailed description	Number of 'other' workers (see section 5.2.3) who were not employed by this establishment in the preceding period.

6.26.6 Total inflows

Variable name	
Origin Detailed description	Computable. Number of new workers in the establishment, i. e., workers who were not employed by this establishment in the preceding period. Can be calculated as: in_nml + in_app + in_pr + in_mpt + in_other.

6.27 Package 403: age

6.27.1 Outflows aged 15–17 years (out _age_1)

Variable name	out_age_1
Origin Detailed description	Outflows by age group. Age is calculated with respect to the end of the period.

6.27.2 Outflows aged 18–24 years (out _age_2)

Variable name	out_age_2
Origin Detailed description	Outflows by age group. Age is calculated with respect to the end of the period.

6.27.3 Outflows aged 25–29 years (out _age_3)

Variable name	out_age_3
Origin Detailed description	Outflows by age group. Age is calculated with respect to the end of the period.

6.27.4 Outflows aged 30–44 years (out _age_4)

Variable name	out_age_4
Origin Detailed description	Outflows by age group. Age is calculated with respect to the end of the period.

6.27.5 Outflows aged 45–49 years (out _age_5)

Variable name	out_age_5
Origin Detailed description	Outflows by age group. Age is calculated with respect to the end of the period.

6.27.6 Outflows aged 50–54 years (out _age_6)

Variable name	out_age_6
Origin Detailed description	Outflows by age group. Age is calculated with respect to the end of the period.

6.27.7 Outflows aged 55–59 years (out _age_7)

Variable name	out_age_7
Origin Detailed description	Outflows by age group. Age is calculated with respect to the end of the period.

6.27.8 Outflows aged 60 years and older (out_age_8)

Variable name	out_age_8
Origin Detailed description	Outflows by age group. Age is calculated with respect to the end of the period.

6.28 Package 404: tenure

6.28.1 Outflows after 1 quarter of job tenure (out_ten_1)

Variable name	out_ten_1
Origin Detailed description	End-of-period outflows of workers who had a certain tenure within the establishment.

6.28.2 Outflows after 2–4 quarters of job tenure (out_ten_2)

Variable name	out_ten_2
Origin Detailed description	End-of-period outflows of workers who had a certain tenure within the establishment.

6.28.3 Outflows after 1–3 years of job tenure (out_ten_3)

Variable name	out_ten_3
Origin Detailed description	End-of-period outflows of workers who had a certain tenure within the establishment.

6.28.4 Outflows after 4–9 years of job tenure (out_ten_4)

Variable name	out_ten_4
Origin Detailed description	End-of-period outflows of workers who had a certain tenure within the establishment.

6.28.5 Outflows after more than 9 years of job tenure (out_ten_5)

Variable name	out_ten_5
Origin Detailed description	End-of-period outflows of workers who had a certain tenure within the establishment.

6.29 Package 406: worker flows

6.29.1 Outflows using daily count (out_dc)

Variable name	out_dc
Origin Detailed description	Number of “entry-events” (a worker employed today but not the preceding day) in the current period.

6.29.2 Outflows according to LEHD definition (out _lehd)

Variable name	out_lehd
Origin Detailed description	Number of workers employed for at least 1 day in T-1, but not in T.

6.29.3 Outflows according to LEHD full-period definition (out _lehd_fp)

Variable name	in_lehd_fp
Origin Detailed description	Number of workers employed for at least 1 day in each the current and the two preceding periods but not employed in the subsequent period.

6.30 Package 407: sex and hours of work

6.30.1 Outflows using the standard end-of-period definition (out_eop)

Variable name	out_eop
Origin Detailed description	Number of regular workers employed at the end of the preceding period but not employed as regular workers at the end of the current period in the same establishment.

6.30.2 Male outflows (out_male)

Variable name	out_male
Origin Detailed description	Number of male outflows.

6.30.3 Female outflows

Variable name	
Origin Detailed description	Computable. Number of female outflows. Can be calculated as: out_eop - out_male

6.31 Package 408: qualification

6.31.1 Outflows of low-skilled workers (out_qual_1)

Variable name	out_qual_1
Origin Detailed description	 Number of outflows without formal vocational training

6.31.2 Outflows of medium-skilled workers (out_qual_2)

Variable name	out_qual_2
Origin Detailed description	 Number of outflows with formal vocational training

6.31.3 Outflows of high-skilled workers (out_qual_3)

Variable name	out_qual_3
Origin Detailed description	 Number of outflows with an university degree (*Universität* or *Fachhochschule*)

6.32 Package 409a: flows to unemployment (ALG / ALG1)

Note that the definition of the duration of unemployment depends on the frequency of the package.

6.32.1 Outflows to unemployment I (out_ue1)

Variable name	out_ue1
Origin Detailed description	 Number of outflows subsequently unemployed for 1 quarter (yearly data: 1 year)

6.32.2 Outflows to unemployment II (out_ue2)

Variable label Variable name	 out_ue2
Origin Detailed description Sensitive variable	 Number of outflows subsequently unemployed for 2 quarters (2 years)

6.32.3 Outflows to unemployment III (out_ue3)

Variable name	out_ue3
Origin Detailed description	 Number of outflows subsequently unemployed for 3–4 quarters (3 years)

6.32.4 Outflows to unemployment IV (out_ue4)

Variable name	out_ue4
Origin Detailed description	Number of outflows subsequently unemployed for 5–12 quarters (4+ years)

6.32.5 Outflows to unemployment V (out_ue5)

Variable name	
Origin Detailed description	Computable. Number of outflows subsequently unemployed for more than 12 quarters. (Yearly data: out_ue5 ≡ 0)

6.33 Package 409b: flows to non-employment (excl. ALG / ALG1)

Note that the definition of the duration of non-employment depends on the frequency of the package.

6.33.1 Outflows to non-employment I (out_oolf1)

Variable name	out_oolf1
Origin Detailed description	Number of outflows subsequently economically inactive for 1 quarter (1 year)

6.33.2 Outflows to non-employment II (out_oolf2)

Variable name	out_oolf2
Origin Detailed description	Number of outflows subsequently economically inactive for 2 quarters (2 years)

6.33.3 Outflows to non-employment III (out_oolf3)

Variable name	out_oolf3
Origin Detailed description	Number of outflows subsequently economically inactive for 3-4 quarters (3 years)

6.33.4 Outflows to non-employment IV (out_oolf4)

Variable name	out_oolf4
Origin Detailed description	Number of outflows subsequently economically inactive for 5-12 quarters (4+ years)

6.33.5 Outflows to non-employment V (out_oolf5)

Variable name	
Origin Detailed description	Computable. Number of outflows previously economically inactive for more than 12 quarters. (Yearly data: out_oolf5 ≡ 0)

6.34 Package 409c: decomposition of outflows

6.34.1 Outflows to unemployment (out_ue)

Variable name	out_ue
Origin Detailed description	Outflows who were unemployed, i. e., who received unemployment benefits, in the current period.

6.34.2 Outflows to non-employment (out_oolf)

Variable name	out_oolf
Origin Detailed description	Outflows who were neither unemployed nor employed in the current period.

6.34.3 Outflows to employment (out_e)

Variable name	out_e
Origin Detailed description	Outflows who keep being employed (i. e., who changed the employer/establishment)

6.34.4 End of career outflows

Variable name	
Origin Detailed description	Computable. Number of workers regularly employed in the preceding period for the last time and not registered as unemployed afterwards. Can be calculated as: out_end = out_eop - (out_ue + out_oofl + out_e).

6.35 Package 410: wage change

Note: Wages might be right-censored. Employees with censored wages in both observation periods will be classified as workers with rigid wages.

6.35.1 Outflows with wage decrease by at least 2 percent (out_dw_dec2)

Variable name	out_dw_dec2
Origin Detailed description	Number of outflows (to employment) experiencing a nominal wage decrease by at least 2/4 percent.

6.35.2 Outflows with wage decrease by at least 4 percent (out_dw_dec4)

Variable name	out_dw_dec4
Origin Detailed description	Number of outflows (to employment) experiencing a nominal wage decrease by at least 2/4 percent.

6.35.3 Outflows with wage increase by at least 2 percent (out_dw_inc2)

Variable name	out_dw_inc2
Origin Detailed description	Number of outflows (to employment) experiencing a nominal wage increase by at least 2/4 percent.

6.35.4 Outflows with wage increase by at least 4 percent (out_dw_inc4)

Variable name	out_dw_inc4
Origin Detailed description	Number of outflows (to employment) experiencing a nominal wage increase by at least 2/4 percent.

6.35.5 Outflows with an absolute wage change less than 2 percent (out_dw_rig2)

Variable name	out_dw_rig2
Origin Detailed description	Number of outflows (to employment) experiencing nominal wage rigidity, i. e., an absolute wage change less than 2/4 percent.

6.35.6 Outflows with an absolute wage change less than 4 percent (out_dw_rig4)

Variable name	out_dw_rig4
Origin Detailed description	Number of outflows (to employment) experiencing nominal wage rigidity, i. e., an absolute wage change less than 2/4 percent.

6.35.7 Outflows from young establishments with wage decrease by at least 4 percent (out_dw_dec4_yng)

Variable name	out_dw_dec4_yng
Origin Detailed description	Number of outflows to young establishments experiencing a nominal wage decrease by at least 4 percent. Establishments are classified as young when they have been operating for not more than 3 quarters (quarterly frequency) or 1 year (yearly frequency).

6.35.8 Outflows from young establishments with wage increase by at least 4 percent (out_dw_inc4_yng)

Variable name	out_dw_inc4_yng
Origin Detailed description	Number of outflows to young establishments experiencing a nominal wage increase by at least 4 percent. Establishments are classified as young when they have been operating for not more than 3 quarters (quarterly frequency) or 1 year (yearly frequency)

6.35.9 Outflows from young establishments with an absolute wage change less than 4 percent (out_dw_rig4_yng)

Variable name	out_dw_rig4_yng
Origin Detailed description	Number of outflows to young establishments experiencing nominal wage rigidity, i.e., an absolute wage change less than 4 percent. Establishments are classified as young when they have been operating for not more than 3 quarters (quarterly frequency) or 1 year (yearly frequency).

6.36 Package 410_imp: wage change

Same as Package 410 (see section 6.35), but using imputed wages (see section 3.3.3). Hence, might are definitely not right-censored.

6.37 Package 411: qualification according to Blossfeld

For information concerning (the validation of) the occupation data, used to form the categories of the Blossfeld classification (see Blossfeld 1987), please refer to sections 4.6 and 3.1.5 of Schmucker et al. (2016).

6.37.1 Outflows classified as low-skilled according to Blossfeld (out_task_1)

Variable name	out_task_1
Origin Detailed description	Agricultural occupations, elementary manual occupations, elementary personal services occupations, elementary administrative occupations

6.37.2 Outflows classified as medium-skilled according to Blossfeld (out_task_2)

Variable name	out_task_2
Origin Detailed description	Skilled manual occupations, skilled services occupations, skilled administrative occupations

6.37.3 Outflows classified as semi-skilled according to Blossfeld (out_task_3)

Variable name	out_task_3
Origin Detailed description	Technicians, associate professionals

6.37.4 Outflows classified as high-skilled according to Blossfeld (out_task_4)

Variable name	out_task_4
Origin Detailed description	Professional occupations, managers

6.38 Package 412: status change outflows

6.38.1 External outflows

Variable name	
Origin Detailed description	Computable. Number of outflows who continue to work but not as regular workers, and in another establishment. Can be calculated as: out_app_ext + out_mpt_ext + out_pr_ext + out_other_ext + out_pt_ext

6.38.2 Internal outflows

Variable name	
Origin Detailed description	Computable. Number of outflows who stay in the establishment (but no longer employed as regular workers). Can be calculated as: out_app_int + out_mpt_int + out_pr_int + out_other_int + out_pt_int

6.38.3 External outflows of former apprentices (out_app_ext)

Variable name	out_app_ext
Origin Detailed description	Number of outflows employed as apprentices in **another establishment** at the preceding reference date.

6.38.4 External outflows of former marginal part-time workers (out_mpt_ext)

Variable name	out_mpt_ext
Origin Detailed description	Number of outflows employed marginal part-time workers in **another establishment** at the preceding reference date.

6.38.5 External outflows of workers previously in partial retirement (out_pr_ext)

Variable name	out_pr_ext
Origin Detailed description	Number of outflows employed workers in partial retirement in **another establishment** at the preceding reference date.

6.38.6 External outflows of former ‘other’ workers (out_other_ext)

Variable name	out_other_ext
Origin Detailed description	Number of outflows employed ‘other’ workers (see section 5.2.3) in **another establishment** at the preceding reference date.

6.38.7 External outflows of former 'normal' workers (out_nml _ext)

Variable name	out_pt_ext
Origin Detailed description	Number of outflows employed 'normal' workers (see section 5.2.2) in part-time in **another establishment** at the preceding reference date.

6.38.8 Internal outflows of former apprentices (out_app_int)

Variable name	out_app_int
Origin Detailed description	Number of outflows employed as apprentices by **the same establishment** at the preceding reference date.

6.38.9 Internal outflows of former marginal part-time workers (out_mpt_int)

Variable name	out_mpt_int
Origin Detailed description	Number of outflows employed marginal part-time workers by **the same establishment** at the preceding reference date.

6.38.10 Internal outflows of workers previously in partial retirement (out_pr_int)

Variable name	out_pr_int
Origin Detailed description	Number of outflows employed workers in partial retirement by **the same establishment** at the preceding reference date.

6.38.11 Internal outflows of former ‘other’ workers (out_other_int)

Variable name	out_pt_int
Origin Detailed description	Number of outflows employed ‘other’ workers (see section 5.2.3) by **the same establishment** at the preceding reference date.

6.38.12 Internal outflows of former 'normal' workers (out_nml_int)

Variable name	out_pt_int
Origin Detailed description	Number of outflows employed 'normal' workers (see section 5.2.2) in part-time by **the same establishment** at the preceding reference date.

6.39 Package 413: permanent and temporary outflows

6.39.1 Permanent outflows (out_perm)

Variable name	out_perm
Origin Detailed description	Number of workers regularly employed in the preceding period but not employed (any status) in this establishment in the current or any of the 3 subsequent periods.

6.39.2 Temporary outflows (out_temp)

Variable name	out_temp
Origin Detailed description	Number of workers regularly employed in the preceding period, not employed (in this establishment) in the current period, and again employed by the establishment in at least one of the 3 subsequent periods.

6.40 Package 414: apprentices, partial retirement, marginal part-time, and interns (not calculated on a regular worker basis!)

6.40.1 Outflow of normal workers (out_nml)

Variable name	out_nml
Origin Detailed description	Number of 'normal' workers (see section 5.2.2) who left the establishment

6.40.2 Outflow of apprentices (out_app)

Variable name	out_app
Origin Detailed description	Number of apprentices who left the establishment

6.40.3 Outflow of partial retirement workers (out_pr)

Variable name	out_pr
Origin Detailed description	Number of workers in partial retirement who left the establishment

6.40.4 Outflow of marginal part-time workers (out_mpt)

Variable name	out_mpt
Origin Detailed description	Number of marginal part-time workers who left the establishment

6.40.5 Outflow of other workers (out_other)

Variable name	out_other
Origin Detailed description	Number of 'other' workers (see section 5.2.3) who left the establishment

6.40.6 Total outflows

Variable name	
Origin Detailed description	Computable. Total number of workers who left the establishment. Can be calculated as: out_nml + out_app + out_pr + out_mpt + out_other.

6.41 Package 515: stayer (change of person group)

6.41.1 Former ‘normal’ workers (stay_nml)

Variable name	stay_nml
Origin Detailed description	Number of workers employed as 'normal' workers (see section 5.2.2) in the preceding period who are still employed in the same establishment in the current period, but not as 'normal' workers.

6.41.2 Former apprentices (stay_app)

Variable name	stay_app
Origin Detailed description	Number of workers employed as apprentices in the preceding period who are still employed in the same establishment in the current period, but not as apprentices.

6.41.3 Former partially retirees (stay_pr)

Variable name	stay_pr
Origin Detailed description	Number of workers in partial retirement in the preceding period who are still employed in the same establishment in the current period, but not as partially retired workers.

6.41.4 Former marginal part-time workers (stay_mpt)

Variable name	stay_mpt
Origin Detailed description	Number of workers employed as marginal part-time workers in the preceding period who are still employed in the same establishment in the current period, but not as marginal part-time workers.

6.41.5 Former 'other' workers (stay_other)

Variable name	stay_other
Origin Detailed description	Number of workers employed as 'other' workers (see section 5.2.3) in the preceding period who are still employed in the same establishment in the current period, but not as 'other' workers.

6.41.6 Former marginal part-time workers (out_mptw_101)

Variable name	out_mptw_101
Origin Detailed description	Number of workers at the end of the period who were marginal part-time employed at the end of the preceding period.

6.41.7 Former other workers (out_other_101)

Variable name	out_other_101
Origin Detailed description	Number of regular workers at the end of the period who were employed as ‘other’ workers (see section 5.2.3) at the end of the preceding period.

6.41.8 Former regular workers (out_regw_101)

Variable name	out_regw_101
Origin Detailed description	Number of regular workers at the end of the period who were employed as regular workers at the end of the preceding period.

7 References

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

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

Bundesagentur für Arbeit (Ed.) (2007): Handbuch für die Betriebsnummernvergabe und -pflege im Rahmen des Meldeverfahren zur Sozialversicherung. Edition December 2007, Nürnberg.

Bundesagentur für Arbeit (Ed.) (2010): Beschäftigungsstatistik: Umstellung der Klassifikation der Wirtschaftszweige von WZ 2003 auf WZ 2008. Nürnberg.

Bundesanstalt für Arbeit (Ed.) (1973): Verzeichnis der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Nürnberg.

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

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

Cramer, Ulrich (1985): Probleme der Genauigkeit der Beschäftigtenstatistik. Allgemeines Statistisches Archiv, 69, 56–68.

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.

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.

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

Kruppe, Thomas; Matthes, Britta; Unger, Stefanie (2014): Effectiveness of data correction rules in process-produced data: the case of educational attainment. IAB-Discussion Paper, 15/2014.

Ludsteck, Johannes; Thomsen, Ulrich (2016): Imputation of the working time information for the employment register data. FDZ-Methodenreport, 01/2016 (en).

Schmucker, Alexandra; Seth, Stefan; Ludsteck, J.; Eberle, Johanna; Ganzer, Andreas (2016): Establishment History Panel 1975–2014. FDZ-Datenreport, 03/2016 (en).

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

8 Appendix: person group codes in the BeH*

Code	Name
101	Employees subject to social security with no special features
102	Trainees / apprentices with no special features
103	Employees in partial retirement
104	Freelance home workers
105	Interns
106	Student trainees
108	Recipients of early retirement benefit
109	Marginal part-time employees
110	Short-term employees
112	Family workers in agriculture
116	Recipients of compensation according to the Act on Support in the Case of Termination of Farming Activities
118	Casual workers
119	Old-age pensioners exempt from insurance contributions and recipients of old-age pension benefits
120	Persons who are presumed to be in employment
121	Trainees / apprentices (earnings not above the low-wage threshold)
122	Trainees / apprentices (external institution)
123	Persons completing a year of voluntary social or environmental work or Federal Voluntary Service
124	Home workers
140	Seamen
141	Trainees/apprentices in seafaring occupations with no special features
142	Seamen in partial retirement
143	Maritime pilots
144	Trainees/apprentices in seafaring occupations (earnings not above the low-wage threshold)
149	Old-age pensioners exempt from insurance contributions and recipients of old-age pension benefits employed in seafaring occupations
190	Employees who are insured solely in the statutory accident insurance
201	Employees in private households (reported via the “household cheque procedure”)
202	Short-term employees
203	Artists and publicists subject to social security
205	Casual workers
207	Nurses in the sense of § 19 SGB XI/with no eligibility for financial assistance on the part of the person receiving nursing care
208	Nurses in the sense of § 19 SGB XI/with eligibility for financial assistance on the part of the person receiving nursing care
209	Marginal part-time employees in private households (reported via the “household cheque procedure”)
210	Short-term employees in private households (reported via the “household cheque procedure”)
301	Persons performing basic military service or voluntary military service
302	Persons performing reserve duty
303	Persons performing alternative civilian service
304	Persons completing a year of voluntary social or environmental work instead of alternative civilian service
305	Persons performing military service, special types
306	Special assignment abroad
599	Miscellaneous workers
999	No details available
XXX	No allocation possible
YYY	Error in original value
ZZZ	No details reported