

Seth, Stefan; Stüber, Heiko

Working Paper

The Administrative Wage and Labor Market Flow Panel

FAU Discussion Papers in Economics, No. 01/2017

Provided in Cooperation with:

Friedrich-Alexander University Erlangen-Nuremberg, Institute for Economics

Suggested Citation: Seth, Stefan; Stüber, Heiko (2018) : The Administrative Wage and Labor Market Flow Panel, FAU Discussion Papers in Economics, No. 01/2017, Friedrich-Alexander-Universität Erlangen-Nürnberg, Institute for Economics, Nürnberg

This Version is available at:

<https://hdl.handle.net/10419/189829>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

No. 01/2017

**The Administrative Wage and Labor Market
Flow Panel**

Stefan Seth
Institute for Employment Research (IAB)

Heiko Stüber
University of Erlangen-Nürnberg and IAB

ISSN 1867-6707

The Administrative Wage and Labor Market Flow Panel

Heiko Stüber, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Institute for Employment Research (IAB), IZA, and Labor and Socio-Economic Research Center (LASER)

Stefan Seth, Institute for Employment Research (IAB)

Abstract

This paper describes the updated version of the Administrative Wage and Labor Market Flow Panel (AWFP, v1.1). The AWFP is a dataset on labor market flows and stocks for the universe of German establishments covering the years 1975–2014. It contains data on job flows, worker flows, and wages for each establishment. The AWFP contains these information also for partitions of the labor force according to various employee characteristics and for some subgroups of employees. The AWFP is available on the annual and quarterly frequency.

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

AWFP dataset: Version 1.1 (AWFP 1975–2014, v1.1)

Documentation: Version 1.1

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

Preparation of the AWFP dataset: Stefan Seth (IAB)

Aggregated public release AWFP datasets: Version 1.0 (Agg_AWFP 1975–2014, v1.0)

Preparation of the aggregated public release AWFP datasets: Heiko Stüber (FAU and IAB)

Note: For a data report on version 1.0 of the AWFP see Seth and Stüber (2017). For a data report of the aggregated public release AWFP datasets see Stüber and Seth (2017).

Acknowledgements

The basic data generation process of the AWFP is identical to the one of the Establishment History Panel (BHP) 1975–2014. Therefore, some sections of this data report (marked with an asterisk, *) are copied (and only slightly altered) from the data report of the BHP (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

1	Introduction and outline	5
1.1	Introduction	5
1.2	Data access	6
1.3	Outline of the AWF	8
1.4	Differences from the predecessor version	8
2	Data sources	9
2.1	Employee history (BeH) *	9
2.2	Benefit recipient history (LeH)	9
3	Data preparation – correction and validation procedures performed on the micro-level data	10
3.1	Selection of the notifications in the BeH*	10
3.2	Imputation of the education/training data *	10
3.3	Imputation of the information on earnings	11
3.4	Imputation of the information on full-time and part-time employment*	12
3.5	Strike corrections*	12
4	Data quality	13
4.1	Eastern Germany *	13
4.2	Under-recording of notifications in the latest available data *	13
4.3	Data on earnings *	13
4.4	Part-time employees*	14
4.5	Classification of economic activities	14
5	Generating the Administrative Wage and Labor Market Flow Panel	15
5.1	Overview of the dataset	15
5.2	Definitions	16
5.3	Calculation of omitted variables	17
5.4	Merging packages	17
5.5	Programming example	18
6	Description of the variables and characteristics	19
6.1	Common identifiers	19
6.2	Package p_101: dates	19
6.3	Package pa102: location and industry	20
6.4	Package 103: mean age	21
6.5	Package 104: mean tenure	21
6.6	Package 105_x: wages (regular workers, stayers, inflows, outflows)	21
6.7	Package 105imp_x: imputed wage (regular workers, stayers, inflows, outflows)	22
6.8	Package 105_in_y: (imputed) wages for inflows (ue, strt, oolf, ne, e)	23
6.9	Package 203: age	23
6.10	Package 204: tenure	24
6.11	Package 206: workers – various stock definitions	25
6.12	Package 207: standard-definition stock and sex	25
6.13	Package 208: qualification	25

6.14	Package 211: qualification according to Blossfeld	26
6.15	Package 214: apprentices, partial retirement, marginal part-time, and interns (not calculated on a regular worker basis!)	26
6.16	Package 303: age inflows	27
6.17	Package 306: various inflow definitions	28
6.18	Package 307: standard-definition inflows and sex	29
6.19	Package 308: qualification inflows	29
6.20	Package 309a: flows from unemployment (ALG / ALG1)	29
6.21	Package 309b: flows from non-employment (excl. ALG / ALG1)	30
6.22	Package 309c: decomposition of inflows	31
6.23	Package 310: wage change	31
6.24	Package 311: qualification according to Blossfeld	33
6.25	Package 312: status change inflows	33
6.26	Package 313: hires and re-hires	35
6.27	Package 314: apprentices, partial retirement, marginal part-time, and interns (not calculated on a regular worker basis!)	35
6.28	Package 403: age outflows	36
6.29	Package 404: tenure outflows	37
6.30	Package 406: worker flows	38
6.31	Package 407: standard-definition outflows and sex	38
6.32	Package 408: qualification	39
6.33	Package 409a: flows to unemployment (ALG / ALG1)	39
6.34	Package 409b: flows to non-employment (excl. ALG / ALG1)	40
6.35	Package 409c: decomposition of outflows	41
6.36	Package 410: wage change	41
6.37	Package 410_imp: wage change	42
6.38	Package 411: qualification according to Blossfeld	43
6.39	Package 412: status change outflows	43
6.40	Package 413: permanent and temporary outflows	45
6.41	Package 414: apprentices, partial retirement, marginal part-time, and interns (not calculated on a regular worker basis!)	45
6.42	Package 515: stayer (change of person group)	46
6.43	Package pm_eop: monthly frequency (stock and flows)	47
7	References	48
8	Appendix 1: person group codes in the BeH*	50
9	Appendix 2: end of period reference dates	51

Note: Some sections of this data report (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. In addition, the AWFP contains this information 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 on the annual and quarterly frequency¹. Some variables are also available on a monthly basis (see Section 1.4).

The main data source of the AWFP data is the Employment History (*Beschäftigten-Historik*, BeH) of the 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).

This paper describes the updated Administrative Wage and Labor Market Flow Panel (AWFP, v1.1). Differences compared with the predecessor version (AWFP, v1.0; see Seth and Stüber 2017) are specified in Section 1.4.

The Research Data Centre (FDZ) of the German Federal Employment Agency will provide a 50% random sample of the AWFP (see Section 1.2.1) and AWFP extensions for FDZ datasets (see Section 1.2.2). We already provide public release datasets of the AWFP (see Section 1.2.3 and Stüber and Seth (2017)).³

¹ There is one exception: packages 105_y (see Section 6.8) are only available at the quarterly frequency.

² The BeH also comprises marginal part-time workers employed since 1999.

³ In Stüber and Seth (2017) we do not only introduce the parts of the AWFP that are relevant for generating the aggregated public release datasets, but we also give an outlook on forthcoming and planned datasets using the AWFP.

1.2 Data access

1.2.1 FDZ version of the AAFP (FDZ-AAFP)

For a 50% random sample of all establishments, selected AAFP variables will be available at the Research Data Centre (FDZ) of the German Federal Employment Agency on the annual and quarterly frequency. The release date is scheduled for the first half of 2019.⁴ The FDZ-AAFP 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, it is first necessary to submit an application to the FDZ.

The FDZ-AAFP will contain the following variables:

- Index of year / quarter (a / q).
- Establishment identifier (betnr).
- Stock regular workers (st_eop; package 207).
- Stock of regular workers according to qualification (st_qual_1, st_qual_2, st_qual_3; package 208).
- Stock all workers — including part-time, apprentices etc. (st_all; calculated using package 214).
- Mean age of regular workers (mean_age; package 103).
- Mean tenure of regular workers (mean_tenure; package 104).
- Imputed mean daily wages of regular workers (dw_imp_mean; package 105imp).
- Standard deviation of imputed daily wage of regular workers (dw_imp_sd; package 105imp).
- Federal state (state; calculated using package 102).
- Industry classification (w93_imp; package 102).
- Inflows / outflows regular workers (in_eop / out_eop; package 307 / 407).
- Imputed mean daily wages of inflows, stayers, and outflows (dw_imp_mean_in, dw_imp_mean_st, dw_imp_mean_out; packages pq105imp_in, pq105imp_st, pq105imp_out). These variables are only available on the quarterly frequency.
- Inflows / outflows regular workers according to qualification (in_qual_1, in_qual_2, in_qual_3 / out_qual_1, out_qual_2, out_qual_3; package 308 / 408).
- Inflow / outflows all workers — including part-time, apprentices etc. (in_all / out_all; calculated using package 314 / 414).

⁴ In Stüber and Seth (2017), we announced an AAFP extension for the Establishment History Panel (BHP). Since the BHP is only available at an annual frequency as of June 30, we decided that it is more useful for researchers to provide a FDZ-AAFP at a quarterly frequency instead.

- Inflows / outflows of regular workers from / to employment (in_ee / out_ee; package 309c / 409c)
- Newly hired regular workers (in_new; package 313)
- Re-hired regular workers (in_rehire ; package 313)
- Permanent outflows (out_perm; package 413)
- Temporary outflows (out_temp; package 413)

1.2.2 AWFP extensions for FDZ datasets

We are working on AWFP extensions for the IAB Job Vacancy Survey and the IAB Establishment Panel. These extensions will probably be available at the FDZ in 2019. For further details, please refer to Stüber and Seth (2017). The IAB Job Vacancy Survey, the IAB Establishment Panel and the corresponding AWFP extensions may be analyzed in the context of a research visit at the FDZ and subsequent remote data access. In order to use the data, it is necessary to submit an application to the FDZ.

1.2.3 Public release data

The aggregated public release data contain, among other things, information on job and worker flows for groups of firms and can thus be used to study the cyclical dynamics (of fractions) of the labor market in terms of turnover and churning. The aggregated public release data of the AWFP cover the time period 1975–2014.

Thus far, six aggregated public release data sets are available. All data sets are available on the quarterly and on the annual frequency. They are 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/>). For a detailed description of the public release data, please refer to Stüber and Seth (2017).

1.3 Outline of the AAFP

Categories	Descriptions
Topics	<ul style="list-style-type: none"> • 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: December 31 of each year Quarterly frequency: last day of each quarter Monthly frequency: last day of each month
Regional structure	Districts (<i>Kreise</i>)
Type of territorial allocation	Corrected territorial allocation as of December 31, 2014.
Frequency of data collection	Annual frequency Quarterly frequency Monthly frequency
File format	Stata

1.4 Differences from the predecessor version

Compared to the predecessor version (AAFP 1975–2014, v1.0; see Seth and Stüber 2017), the following changes occurred:

1) New data packages:

- a) Further disaggregated wage information are available. Packages `pq105_y` contain the mean (imputed) wages for inflows, disaggregated by their origin (unemployment, employment, etc.). These packages are only available on the quarterly frequency. See section 6.8 for a detailed description.

2) Monthly frequency:

- a. Three variables are now also available on the monthly frequency (see Section 6.43): the stock of regular workers (`st_eop`), the inflow of workers (`in_eop`), and the outflow of workers (`out_eop`).

2 Data sources

2.1 Employee history (BeH)*

The data source 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 January 1, 1973 and was extended to cover eastern Germany as of January 1, 1991 (for further details see Bender et al. 1996 and Wermter/Cramer 1988). Under the so-called DEÜV procedure (previously DEVO/DÜVO), employers are required to submit notifications to the responsible social security agencies concerning all 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 not recorded in the BeH. Employees in marginal part-time employment and unpaid family workers have been contained in the data from April 1, 1999 onwards. Employers are obliged to report the exact start and ending dates of an employment relationship and to yearly confirm an existing one, so that it is possible to track workers' careers on a daily basis. The observation period of the BeH V10.0.0 extends from January 1, 1975 to December 31, 2014.

2.1.1 The establishment concept*

It is important to understand that the BeH uses a specific definition of 'establishment': An establishment is a "regionally and economically delimited unit in which employees work". It 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 which 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 – correction and validation procedures performed on the micro-level data

For compiling the AWF, 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 a number of 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 AWF, 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 1).
- Notifications with a wage of zero 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 Imputation of the education/training data*

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 percent 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). While this actually permits a more precise recording of education and training qualifications, no time-consistent information is available for the entire period. In order to achieve that, the methods of recording the data are being made compatible by assigning to every combination of values from the new code the most suitable details on education and training according to the old code. However, this has no effect on missing values. So, 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 Imputation 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 AWF.

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 has been 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 percent 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 percent 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 approximately 10 percent of the information on full-time employees' earnings is censored. This leads to biased estimation results as 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 percent) 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 mitigate these issues, the information on earnings (average daily wages) were imputed before the statistics (means and medians) were calculated. The procedure was implemented following Card et al. (2015) and is explained in more detail in Section 8.2 of Schmucker et al. (2016).

3.4 Imputation 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 and working time in the social security notifications. In a good 10 percent of the notifications submitted by the establishments between December 1, 2011 and May 31, 2012, the information regarding working hours is therefore missing. For this reason, a logit model has been developed at the IAB which can be used to impute the missing information (see Ludsteck and Thomsen 2016).

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. These gaps frequently include the reference date of June 30, 1984, which is relevant for the AWFEP on the quarterly frequency. Leaving these gaps in the data would result in considerable distortions in the AWFEP for the industries in the federal states affected in 1984. These gaps have therefore been filled in accordance with the following heuristics:

To identify gaps resulting from lockouts the following definition was used. An account is regarded as locked out if:

- there was a notification in Baden-Wuerttemberg or Hesse on April 30, 1984 that was classified as belonging to the economic activity 280 or 281 (notification 1),
- there was a further notification from the same establishment in July 1984 (notification 2) and
- there was a gap in employment lasting more than five days in May or June 1984.

These gaps have been 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 have been added accordingly.

4 Data quality

The data quality of the AAFP 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 April 15 (or by February 15 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 percent degree of completeness by definition. To generate the AAFP data of 2012, a 30-month file was used, 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. Comparing the 6-, 12- and 18-month files for 2013 one can observe that the 12-month file contains 0.8 percent 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 percent 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 percent between the 12-month and the 18-month files, the increase in the number of establishments is only 0.5 percent. During this period more establishments with more than 200 employees were added to the data.

4.3 Data on earnings*

In 1984 the employment notification procedure changed. From that time onwards one-off payments were reported as part of the annual earnings, which has led 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 AWF, the classification of economic activities 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 AWF 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

5.1 Overview of the dataset

The Administrative Wage and Labor Market Flow Panel (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.⁵

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

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 if all package variables (excluding establishment and time identifier) are zero or missing. Therefore missing values that will be generated when 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 on 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.

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 The only exception is the data package containing information on the monthly frequency. It contains stock and flow information.

5.2 Definitions

All data packages — except package p_101 — contain information on the annual (a), quarterly (q), or monthly (m) frequency. Hence, when we talk about a “period”, we think of a year, a quarter, or a month.

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 Appendix 1 for more details on the person group in the BeH. The stocks and flows in the AAFP 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” packages (312, 412) inform 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 AWF, some variables that 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 of these variables, which can be found in Section 6, are colored grey.

5.4 Merging packages

When merging packages take the following particularities of the AWF 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 replaced 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 AAFP (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 zeros (0):
replace in_eop = 0 if missing(in_eop) // inflows
replace in_male = 0 if missing(in_male) // male inflows
replace out_eop = 0 if missing(out_eop) // outflows
replace out_male = 0 if missing(out_male) // male outflows
```

Example 2: The following Stata programming example shows how to 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). Appendix 2 shows the end-of-period reference dates for all years, quarters, and month.

6.1.1 Establishment identifier

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

6.1.2 Index of year

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

6.1.3 Index of quarter

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

6.1.4 Index of month

Variable name	m
Detailed description	Starts with 1, with month no 1 being January of 1975.

6.2 Package p_101: dates

Note that the first and last appearance of an establishment number offers a first indication for the times when the establishment was founded or 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

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

6.2.2 First regular worker

Variable name	first_rw_m
Detailed description	The first month in which the establishment has employed a ‘regular worker’; the value is missing if there has not been such a month; possibly left-censored.

6.2.3 Shut down date

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

6.2.4 Last regular worker

Variable name	last_rw_m
Detailed description	The last month in which the establishment employed a ‘regular worker’ (as defined in Section 5.2.1); missing values occur; possibly 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

Variable name	district
Detailed description	The district (<i>Kreis</i>) the establishment is located.

6.3.2 Establishment’s industry classification

Variable name	w73
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

Variable name	w93
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

Variable name	w03
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

Variable name	w08
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

Variable name	w73_imp
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

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

6.4 Package 103: mean age

6.4.1 Mean age

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

6.5 Package 104: mean tenure

6.5.1 Mean tenure

Variable name	mean_tenure
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_x: wages (regular workers, stayers, inflows, outflows)

Notes: x indicates the underlying universe of the package: 105_all = all regular workers; 105_in = new regular workers (inflows); 105_out = outgoing regular workers (outflows); 105_st = incumbent regular workers (stayers). All wage information might be right-censored.

6.6.1 Number of observations

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

6.6.2 Mean wage of workers

Variable name	dw_mean
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

Variable name	dw_sd
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

Variable name	dw_p25
Detailed description	25 th 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

Variable name	dw_p50
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

Variable name	dw_p75
Detailed description	75 th 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_x: imputed wage (regular workers, stayers, inflows, outflows)

This package is constructed like package 105_x (see Section 6.6), except that we have used imputed wages (see Section 3.3.3). Hence, no right-censoring occurs. x indicates the underlying universe of the package: 105imp_all = all regular workers; 105imp_in = new regular workers (inflows); 105imp_out = outgoing regular workers (outflows); 105imp_st = incumbent regular workers (stayers).

6.8 Package 105_in_y: (imputed) wages for inflows (ue, strt, oolf, ne, e)

Notes: y indicates the underlying universe of the respective package. 105_in_ue = inflows from unemployment; 105_in_strt = start of career inflows; 105_in_oolf = inflows from non-employment; 105_in_ne = start of career inflows and inflows from non-employment; 105_in_e = inflows from employment. This package is only available on the quarterly frequency! Non-imputed wage information might be right-censored. See also package 309c (Section 6.22), this package includes the number of underlying workers.

6.8.1 Mean wage of workers

Variable name	dw_mean
Detailed description	Mean daily wage of underlying workers at the end of the period. Wages of outflows are calculated with respect to the preceding period.

6.8.2 Mean imputed wage of workers

Variable name	dw_imp_mean
Detailed description	Mean imputed daily wage (see Section 3.3.3) of underlying workers at the end of the period. Wages of outflows are calculated with respect to the preceding period.

6.9 Package 203: age

6.9.1 Stock of workers aged 15–17

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

6.9.2 Stock of workers aged 18–24

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

6.9.3 Stock of workers aged 25–29

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

6.9.4 Stock of workers aged 30–44

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

6.9.5 Stock of workers aged 45–49

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

6.9.6 Stock of workers aged 50–54

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

6.9.7 Stock of workers aged 55–59

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

6.9.8 Stock of workers aged 60 and older

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

6.10 Package 204: tenure

6.10.1 Stock of workers with a job tenure of up to 1 quarter

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

6.10.2 Stock of workers with a job tenure of 2–4 quarters

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

6.10.3 Stock of workers with a job tenure of 5 quarters – 3 years

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

6.10.4 Stock of workers with a job tenure of 13 quarters – 9 years

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

6.10.5 Stock of workers with a job tenure of more than 9 years

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

6.11 Package 206: workers – various stock definitions

6.11.1 Stock of workers at the beginning of the period

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

Note: st_bop is not included in the annual frequency dataset, because it is identical to st_bop for the first quarter of the corresponding year, provided in the quarterly frequency dataset.

6.11.2 Stock of workers according to LEHD flow definition

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

6.11.3 Stock of workers according to LEHD full-period definition

Variable name	st_lehd_fp
Detailed description	Number of workers employed for at least one day in the current period (t), at least one day in the preceding period ($t-1$), and at least one day in the subsequent period ($t+1$).

6.12 Package 207: standard-definition stock and sex

6.12.1 Stock of workers

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

6.12.2 Stock of male workers

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

6.12.3 Stock of female workers

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

6.13 Package 208: qualification

6.13.1 Stock of low-skilled workers

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

6.13.2 Stock of medium-skilled workers

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

6.13.3 Stock of high-skilled workers

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

6.14 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).

6.14.1 Stock of workers classified as low-skilled according to Blossfeld

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

6.14.2 Stock of workers classified as medium-skilled according to Blossfeld

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

6.14.3 Stock of workers classified as semi-skilled according to Blossfeld

Variable name	st_task_3
Detailed description	Technicians and associate professionals.

6.14.4 Stock of workers classified as high-skilled according to Blossfeld

Variable name	st_task_4
Detailed description	Professional occupations and managers.

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

6.15.1 Stock of normal workers

Variable name	st_nml
Detailed description	Number of all 'normal' workers (see Section 5.2.2).

6.15.2 Stock of apprentices

Variable name	st_app
Detailed description	Number of apprentices/trainees (<i>Auszubildende</i>).

6.15.3 Stock of workers in partial retirement

Variable name	st_pr
Detailed description	Number of workers in partial/progressive retirement (<i>Altersteilzeit</i>).

6.15.4 Stock of marginal part-time workers

Variable name	st_mpt
Detailed description	Number of marginal part-time workers (<i>geringfügig entlohnte Beschäftigte</i>).

6.15.5 Stock of other worker

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

6.15.6 Stock of all employees

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

6.16 Package 303: age inflows

6.16.1 Inflows aged 15–17 years

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

6.16.2 Inflows aged 18–24 years

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

6.16.3 Inflows aged 25–29 years

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

6.16.4 Inflows aged 30–44 years

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

6.16.5 Inflows aged 45–49 years

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

6.16.6 Inflows aged 50–54 years

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

6.16.7 Inflows aged 55–59 years

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

6.16.8 Inflows aged 60 years and older

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

6.17 Package 306: various inflow definitions

6.17.1 Inflows using daily count

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

6.17.2 Inflows according to LEHD definition

Variable name	in_lehd
Detailed description	Number of workers regularly employed for at least one day in the current period (t) but not in the preceding period ($t-1$).

6.17.3 Inflows according to LEHD full-period definition

Variable name	in_lehd_fp
Detailed description	Number of workers regularly employed for at least one day in the current (t), subsequent ($t+1$) and preceding period ($t-1$) but not in the period before ($t-2$).

6.18 Package 307: standard-definition inflows and sex

6.18.1 Inflows using the standard end-of-period definition

Variable name	in_eop
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.

6.18.2 Male inflows

Variable name	in_male
Detailed description	Number of male inflows.

6.18.3 Female inflows

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

6.19 Package 308: qualification inflows

6.19.1 Inflows of low-skilled workers

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

6.19.2 Inflows of medium-skilled workers

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

6.19.3 Inflows of high-skilled workers

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

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

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

6.20.1 Inflows from unemployment I

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

6.20.2 Inflows from unemployment II

Variable name	in_ue2
Detailed description	Number of inflows previously unemployed for two quarters (two years).

6.20.3 Inflows from unemployment III

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

6.20.4 Inflows from unemployment IV

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

6.20.5 Inflows from unemployment V

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

6.21 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.

6.21.1 Inflows from non-employment I

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

6.21.2 Inflows from non-employment II

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

6.21.3 Inflows from non-employment III

Variable name	in_oolf3
Detailed description	Number of inflows previously economically inactive for 3–4 quarters (three years).

6.21.4 Inflows from non-employment IV

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

6.21.5 Inflows from non-employment V

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

6.22 Package 309c: decomposition of inflows

6.22.1 Inflows from unemployment

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

6.22.2 Inflows from non-employment

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

6.22.3 Inflows from employment

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

6.22.4 Start of career inflows

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

6.23 Package 310: wage change

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

6.23.1 Inflows with wage decrease by at least 2 percent

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

6.23.2 Inflows with wage decrease by at least 4 percent

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

6.23.3 Inflows with wage increase by at least 2 percent

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

6.23.4 Inflows with wage increase by at least 4 percent

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

6.23.5 Inflows with an absolute wage change less than 2 percent

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

6.23.6 Inflows with an absolute wage change less than 4 percent

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

6.23.7 Inflows from young establishments with wage decrease by at least 4 percent

Variable name	in_dw_dec4_yng
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 three quarters (quarterly frequency) or one year (yearly frequency).

6.23.8 Inflows from young establishments with wage increase by at least 4 percent

Variable name	in_dw_inc4_yng
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 three quarters (quarterly frequency) or one year (yearly frequency).

6.23.9 Inflows from young establishments with an absolute wage change less than 4 percent

Variable name	in_dw_rig4_yng
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 three quarters (quarterly frequency) or one year (yearly frequency).

6.24 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).

6.24.1 Inflows classified as low-skilled according to Blossfeld

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

6.24.2 Inflows classified as medium-skilled according to Blossfeld

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

6.24.3 Inflows classified as semi-skilled according to Blossfeld

Variable name	in_task_3
Detailed description	Technicians and associate professionals.

6.24.4 Inflows classified as high-skilled according to Blossfeld

Variable name	in_task_4
Detailed description	Professional occupations and managers.

6.25 Package 312: status change inflows

6.25.1 External inflows

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

6.25.2 Internal inflows

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

6.25.3 External inflows of former apprentices

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

6.25.4 External inflows of former marginal part-time workers

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

6.25.5 External inflows of workers previously in partial retirement

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

6.25.6 External inflows of former other workers

Variable name	in_other_ext
Detailed description	Number of inflows employed as 'other' workers (see Section 5.2.3) in another establishment at the preceding reference date.

6.25.7 External inflows of former normal workers in part-time

Variable name	in_pt_ext
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.

6.25.8 Internal inflows of former apprentices

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

6.25.9 Internal inflows of former marginal part-time workers

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

6.25.10 Internal inflows of workers previously in partial retirement

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

6.25.11 Internal inflows of former other workers

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

6.25.12 Internal inflows of former normal workers in part-time

Variable name	in_pt_int
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.

6.26 Package 313: hires and re-hires

6.26.1 New hires

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

6.26.2 Re-hires

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

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

6.27.1 Inflow of normal workers

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

6.27.2 Inflows of apprentices

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

6.27.3 Inflow of partial retirement workers

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

6.27.4 Inflow of marginal part-time workers

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

6.27.5 Inflow of other workers

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

6.27.6 Total inflows

Variable name	
Origin	Computable.
Detailed description	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 follows: $in_nml + in_app + in_pr + in_mpt + in_other$.

6.28 Package 403: age outflows

6.28.1 Outflows aged 15–17 years

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

6.28.2 Outflows aged 18–24 years

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

6.28.3 Outflows aged 25–29 years

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

6.28.4 Outflows aged 30–44 years

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

6.28.5 Outflows aged 45–49 years

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

6.28.6 Outflows aged 50–54 years

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

6.28.7 Outflows aged 55–59 years

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

6.28.8 Outflows aged 60 years and older

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

6.29 Package 404: tenure outflows

6.29.1 Outflows after 1 quarter of job tenure

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

6.29.2 Outflows after 2–4 quarters of job tenure

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

6.29.3 Outflows after 1–3 years of job tenure

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

6.29.4 Outflows after 4–9 years of job tenure

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

6.29.5 Outflows after more than 9 years of job tenure

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

6.30 Package 406: worker flows

6.30.1 Outflows using daily count

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

6.30.2 Outflows according to LEHD definition

Variable name	out_lehd
Detailed description	Number of workers employed for at least one day in period $t-1$, but not in t .

6.30.3 Outflows according to LEHD full-period definition

Variable name	in_lehd_fp
Detailed description	Number of workers employed for at least 1 day in periods t , $t-1$, $t-2$, but not employed in period $t+1$.

6.31 Package 407: standard-definition outflows and sex

6.31.1 Outflows using the standard end-of-period definition

Variable name	out_eop
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.

6.31.2 Male outflows

Variable name	out_male
Detailed description	Number of male outflows.

6.31.3 Female outflows

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

6.32 Package 408: qualification

6.32.1 Outflows of low-skilled workers

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

6.32.2 Outflows of medium-skilled workers

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

6.32.3 Outflows of high-skilled workers

Variable name	out_qual_3
Detailed description	Number of outflows with a university degree (<i>Universität</i> or <i>Fachhochschule</i>).

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

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

6.33.1 Outflows to unemployment I

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

6.33.2 Outflows to unemployment II

Variable name	out_ue2
Detailed description	Number of outflows subsequently unemployed for two quarters (two years).

6.33.3 Outflows to unemployment III

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

6.33.4 Outflows to unemployment IV

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

6.33.5 Outflows to unemployment V

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

6.34 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.

6.34.1 Outflows to non-employment I

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

6.34.2 Outflows to non-employment II

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

6.34.3 Outflows to non-employment III

Variable name	out_oolf3
Detailed description	Number of outflows subsequently economically inactive for 3–4 quarters (three years).

6.34.4 Outflows to non-employment IV

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

6.34.5 Outflows to non-employment V

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

6.35 Package 409c: decomposition of outflows

6.35.1 Outflows to unemployment

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

6.35.2 Outflows to non-employment

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

6.35.3 Outflows to employment

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

6.35.4 End of career outflows

Variable name	
Origin	Computable.
Detailed description	Number of workers regularly employed in the preceding period for the last time and not registered as unemployed afterwards. Can be calculated as follows: $out_{eop} - (out_{ue} + out_{oofl} + out_e)$.

6.36 Package 410: wage change

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

6.36.1 Outflows with wage decrease by at least 2 percent

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

6.36.2 Outflows with wage decrease by at least 4 percent

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

6.36.3 Outflows with wage increase by at least 2 percent

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

6.36.4 Outflows with wage increase by at least 4 percent

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

6.36.5 Outflows with an absolute wage change less than 2 percent

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

6.36.6 Outflows with an absolute wage change less than 4 percent

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

6.36.7 Outflows from young establishments with wage decrease by at least 4 percent

Variable name	out_dw_dec4_yng
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 three quarters (quarterly frequency) or one year (yearly frequency).

6.36.8 Outflows from young establishments with wage increase by at least 4 percent

Variable name	out_dw_inc4_yng
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 three quarters (quarterly frequency) or one year (yearly frequency).

6.36.9 Outflows from young establishments with an absolute wage change less than 4 percent

Variable name	out_dw_rig4_yng
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 three quarters (quarterly frequency) or one year (yearly frequency).

6.37 Package 410_imp: wage change

Same as Package 410 (see Section 6.36), but using imputed wages (see Section 3.3.3). Hence, no right-censoring occurs.

6.38 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).

6.38.1 Outflows classified as low-skilled according to Blossfeld

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

6.38.2 Outflows classified as medium-skilled according to Blossfeld

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

6.38.3 Outflows classified as semi-skilled according to Blossfeld

Variable name	out_task_3
Detailed description	Technicians and associate professionals.

6.38.4 Outflows classified as high-skilled according to Blossfeld

Variable name	out_task_4
Detailed description	Professional occupations and managers.

6.39 Package 412: status change outflows

Note: workers counted as “internal outflows” **stay** in the establishment. They cease to be ‘regular workers’, however.

6.39.1 External outflows

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

6.39.2 Internal outflows

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

6.39.3 External outflows: apprentices

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

6.39.4 External outflows: marginal part-time workers

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

6.39.5 External outflows: partial retirement

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

6.39.6 External outflows: other workers

Variable name	out_other_ext
Detailed description	Number of outflows employed as 'other' workers (see Section 5.2.3) in another establishment at the reference date.

6.39.7 External outflows: part-time workers

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

6.39.8 Internal outflows: apprentices

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

6.39.9 Internal outflows: marginal part-time workers

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

6.39.10 Internal outflows: partial retirement

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

6.39.11 Internal outflows: other workers

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

6.39.12 Internal outflows: normal workers in part-time

Variable name	out_pt_int
Detailed description	Number of outflows employed as ‘normal’ workers in part-time (see Section 5.2.2) by the same establishment at the reference date.

6.40 Package 413: permanent and temporary outflows

6.40.1 Permanent outflows

Variable name	out_perm
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.

6.40.2 Temporary outflows

Variable name	out_temp
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 three subsequent periods.

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

6.41.1 Outflow of normal workers

Variable name	out_nml
Detailed description	Number of ‘normal’ workers (see Section 5.2.2) who left the establishment.

6.41.2 Outflow of apprentices

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

6.41.3 Outflow of partial retirement workers

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

6.41.4 Outflow of marginal part-time workers

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

6.41.5 Outflow of other workers

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

6.41.6 Total outflows

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

6.42 Package 515: stayer (change of person group)

6.42.1 Former normal workers

Variable name	stay_nml
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.

6.42.2 Former apprentices

Variable name	stay_app
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.

6.42.3 Former partial retirees

Variable name	stay_pr
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.

6.42.4 Former marginal part-time workers

Variable name	stay_mpt
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.

6.42.5 Former other workers

Variable name	stay_other
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.

6.43 Package pm_eop: monthly frequency (stock and flows)

The package pm_eop contains stock, inflow and outflow information on the monthly frequency. Besides the three stock and flow variables (see below), it contains the establishment identifier (betnr) and the index of month (m). The package — like most packages of the AWFPP — is based on the regular worker basis (see Section 5.2.1).

Please be aware that inflows and outflows are left censored in January 1975 ($m = 1$) and January 1992 ($m = 205$) for East Germany. Since the BeH data for eastern Germany can only be assumed to be sufficiently complete from 1993 onwards, analyses of eastern German establishments should not begin before 1993. Therefore we suggest to drop all observations for $m=1$ and all observations for East Germany for $m \leq 217$ (January 1993).

6.43.1 Stock of workers

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

6.43.2 Inflows using the standard end-of-period definition

Variable name	in_eop
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); possibly left-censored.

6.43.3 Outflows using the standard end-of-period definition

Variable name	out_eop
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); possibly left-censored.

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.)** (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).
- Seth, Stefan; Stüber, Heiko** (2017): Administrative Wage and Labor Market Flow Panel (AWFP) 1975–2014 (v1.0). FAU Discussion Papers in Economics, 1/2017.

Stüber, Heiko; Seth, Stefan (2017): The Public Release Data of the Administrative Wage and Labor Market Flow Panel. *Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics*, forthcoming. (Published Online: 2017-05-09 | DOI: <https://doi.org/10.1515/jbnst-2017-1002>)

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 1: 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

9 Appendix 2: end of period reference dates

Reference date	Month (m)	Quarter (q)	Annual (a)	Reference date	Month (m)	Quarter (q)	Annual (a)
31-Jan-75	1			31-Jan-95	241		
28-Feb-75	2			28-Feb-95	242		
31-Mar-75	3	1		31-Mar-95	243	81	
30-Apr-75	4			30-Apr-95	244		
31-May-75	5			31-May-95	245		
30-Jun-75	6	2		30-Jun-95	246	82	
31-Jul-75	7			31-Jul-95	247		
31-Aug-75	8			31-Aug-95	248		
30-Sep-75	9	3		30-Sep-95	249	83	
31-Oct-75	10			31-Oct-95	250		
30-Nov-75	11			30-Nov-95	251		
31-Dec-75	12	4	1	31-Dec-95	252	84	21
31-Jan-76	13			31-Jan-96	253		
29-Feb-76	14			29-Feb-96	254		
31-Mar-76	15	5		31-Mar-96	255	85	
30-Apr-76	16			30-Apr-96	256		
31-May-76	17			31-May-96	257		
30-Jun-76	18	6		30-Jun-96	258	86	
31-Jul-76	19			31-Jul-96	259		
31-Aug-76	20			31-Aug-96	260		
30-Sep-76	21	7		30-Sep-96	261	87	
31-Oct-76	22			31-Oct-96	262		
30-Nov-76	23			30-Nov-96	263		
31-Dec-76	24	8	2	31-Dec-96	264	88	22
31-Jan-77	25			31-Jan-97	265		
28-Feb-77	26			28-Feb-97	266		
31-Mar-77	27	9		31-Mar-97	267	89	
30-Apr-77	28			30-Apr-97	268		
31-May-77	29			31-May-97	269		
30-Jun-77	30	10		30-Jun-97	270	90	
31-Jul-77	31			31-Jul-97	271		
31-Aug-77	32			31-Aug-97	272		
30-Sep-77	33	11		30-Sep-97	273	91	
31-Oct-77	34			31-Oct-97	274		
30-Nov-77	35			30-Nov-97	275		
31-Dec-77	36	12	3	31-Dec-97	276	92	23
31-Jan-78	37			31-Jan-98	277		
28-Feb-78	38			28-Feb-98	278		
31-Mar-78	39	13		31-Mar-98	279	93	
30-Apr-78	40			30-Apr-98	280		
31-May-78	41			31-May-98	281		
30-Jun-78	42	14		30-Jun-98	282	94	

Reference date	Month (m)	Quarter (q)	Annual (a)	Reference date	Month (m)	Quarter (q)	Annual (a)
31-Jul-78	43			31-Jul-98	283		
31-Aug-78	44			31-Aug-98	284		
30-Sep-78	45	15		30-Sep-98	285	95	
31-Oct-78	46			31-Oct-98	286		
30-Nov-78	47			30-Nov-98	287		
31-Dec-78	48	16	4	31-Dec-98	288	96	24
31-Jan-79	49			31-Jan-99	289		
28-Feb-79	50			28-Feb-99	290		
31-Mar-79	51	17		31-Mar-99	291	97	
30-Apr-79	52			30-Apr-99	292		
31-May-79	53			31-May-99	293		
30-Jun-79	54	18		30-Jun-99	294	98	
31-Jul-79	55			31-Jul-99	295		
31-Aug-79	56			31-Aug-99	296		
30-Sep-79	57	19		30-Sep-99	297	99	
31-Oct-79	58			31-Oct-99	298		
30-Nov-79	59			30-Nov-99	299		
31-Dec-79	60	20	5	31-Dec-99	300	100	25
31-Jan-80	61			31-Jan-00	301		
29-Feb-80	62			29-Feb-00	302		
31-Mar-80	63	21		31-Mar-00	303	101	
30-Apr-80	64			30-Apr-00	304		
31-May-80	65			31-May-00	305		
30-Jun-80	66	22		30-Jun-00	306	102	
31-Jul-80	67			31-Jul-00	307		
31-Aug-80	68			31-Aug-00	308		
30-Sep-80	69	23		30-Sep-00	309	103	
31-Oct-80	70			31-Oct-00	310		
30-Nov-80	71			30-Nov-00	311		
31-Dec-80	72	24	6	31-Dec-00	312	104	26
31-Jan-81	73			31-Jan-01	313		
28-Feb-81	74			28-Feb-01	314		
31-Mar-81	75	25		31-Mar-01	315	105	
30-Apr-81	76			30-Apr-01	316		
31-May-81	77			31-May-01	317		
30-Jun-81	78	26		30-Jun-01	318	106	
31-Jul-81	79			31-Jul-01	319		
31-Aug-81	80			31-Aug-01	320		
30-Sep-81	81	27		30-Sep-01	321	107	
31-Oct-81	82			31-Oct-01	322		
30-Nov-81	83			30-Nov-01	323		
31-Dec-81	84	28	7	31-Dec-01	324	108	27
31-Jan-82	85			31-Jan-02	325		
28-Feb-82	86			28-Feb-02	326		

Reference date	Month (m)	Quarter (q)	Annual (a)	Reference date	Month (m)	Quarter (q)	Annual (a)
31-Mar-82	87	29		31-Mar-02	327	109	
30-Apr-82	88			30-Apr-02	328		
31-May-82	89			31-May-02	329		
30-Jun-82	90	30		30-Jun-02	330	110	
31-Jul-82	91			31-Jul-02	331		
31-Aug-82	92			31-Aug-02	332		
30-Sep-82	93	31		30-Sep-02	333	111	
31-Oct-82	94			31-Oct-02	334		
30-Nov-82	95			30-Nov-02	335		
31-Dec-82	96	32	8	31-Dec-02	336	112	28
31-Jan-83	97			31-Jan-03	337		
28-Feb-83	98			28-Feb-03	338		
31-Mar-83	99	33		31-Mar-03	339	113	
30-Apr-83	100			30-Apr-03	340		
31-May-83	101			31-May-03	341		
30-Jun-83	102	34		30-Jun-03	342	114	
31-Jul-83	103			31-Jul-03	343		
31-Aug-83	104			31-Aug-03	344		
30-Sep-83	105	35		30-Sep-03	345	115	
31-Oct-83	106			31-Oct-03	346		
30-Nov-83	107			30-Nov-03	347		
31-Dec-83	108	36	9	31-Dec-03	348	116	29
31-Jan-84	109			31-Jan-04	349		
29-Feb-84	110			29-Feb-04	350		
31-Mar-84	111	37		31-Mar-04	351	117	
30-Apr-84	112			30-Apr-04	352		
31-May-84	113			31-May-04	353		
30-Jun-84	114	38		30-Jun-04	354	118	
31-Jul-84	115			31-Jul-04	355		
31-Aug-84	116			31-Aug-04	356		
30-Sep-84	117	39		30-Sep-04	357	119	
31-Oct-84	118			31-Oct-04	358		
30-Nov-84	119			30-Nov-04	359		
31-Dec-84	120	40	10	31-Dec-04	360	120	30
31-Jan-85	121			31-Jan-05	361		
28-Feb-85	122			28-Feb-05	362		
31-Mar-85	123	41		31-Mar-05	363	121	
30-Apr-85	124			30-Apr-05	364		
31-May-85	125			31-May-05	365		
30-Jun-85	126	42		30-Jun-05	366	122	
31-Jul-85	127			31-Jul-05	367		
31-Aug-85	128			31-Aug-05	368		
30-Sep-85	129	43		30-Sep-05	369	123	
31-Oct-85	130			31-Oct-05	370		

Reference date	Month (m)	Quarter (q)	Annual (a)	Reference date	Month (m)	Quarter (q)	Annual (a)
30-Nov-85	131			30-Nov-05	371		
31-Dec-85	132	44	11	31-Dec-05	372	124	31
31-Jan-86	133			31-Jan-06	373		
28-Feb-86	134			28-Feb-06	374		
31-Mar-86	135	45		31-Mar-06	375	125	
30-Apr-86	136			30-Apr-06	376		
31-May-86	137			31-May-06	377		
30-Jun-86	138	46		30-Jun-06	378	126	
31-Jul-86	139			31-Jul-06	379		
31-Aug-86	140			31-Aug-06	380		
30-Sep-86	141	47		30-Sep-06	381	127	
31-Oct-86	142			31-Oct-06	382		
30-Nov-86	143			30-Nov-06	383		
31-Dec-86	144	48	12	31-Dec-06	384	128	32
31-Jan-87	145			31-Jan-07	385		
28-Feb-87	146			28-Feb-07	386		
31-Mar-87	147	49		31-Mar-07	387	129	
30-Apr-87	148			30-Apr-07	388		
31-May-87	149			31-May-07	389		
30-Jun-87	150	50		30-Jun-07	390	130	
31-Jul-87	151			31-Jul-07	391		
31-Aug-87	152			31-Aug-07	392		
30-Sep-87	153	51		30-Sep-07	393	131	
31-Oct-87	154			31-Oct-07	394		
30-Nov-87	155			30-Nov-07	395		
31-Dec-87	156	52	13	31-Dec-07	396	132	33
31-Jan-88	157			31-Jan-08	397		
29-Feb-88	158			29-Feb-08	398		
31-Mar-88	159	53		31-Mar-08	399	133	
30-Apr-88	160			30-Apr-08	400		
31-May-88	161			31-May-08	401		
30-Jun-88	162	54		30-Jun-08	402	134	
31-Jul-88	163			31-Jul-08	403		
31-Aug-88	164			31-Aug-08	404		
30-Sep-88	165	55		30-Sep-08	405	135	
31-Oct-88	166			31-Oct-08	406		
30-Nov-88	167			30-Nov-08	407		
31-Dec-88	168	56	14	31-Dec-08	408	136	34
31-Jan-89	169			31-Jan-09	409		
28-Feb-89	170			28-Feb-09	410		
31-Mar-89	171	57		31-Mar-09	411	137	
30-Apr-89	172			30-Apr-09	412		
31-May-89	173			31-May-09	413		
30-Jun-89	174	58		30-Jun-09	414	138	

Reference date	Month (m)	Quarter (q)	Annual (a)	Reference date	Month (m)	Quarter (q)	Annual (a)
31-Jul-89	175			31-Jul-09	415		
31-Aug-89	176			31-Aug-09	416		
30-Sep-89	177	59		30-Sep-09	417	139	
31-Oct-89	178			31-Oct-09	418		
30-Nov-89	179			30-Nov-09	419		
31-Dec-89	180	60	15	31-Dec-09	420	140	35
31-Jan-90	181			31-Jan-10	421		
28-Feb-90	182			28-Feb-10	422		
31-Mar-90	183	61		31-Mar-10	423	141	
30-Apr-90	184			30-Apr-10	424		
31-May-90	185			31-May-10	425		
30-Jun-90	186	62		30-Jun-10	426	142	
31-Jul-90	187			31-Jul-10	427		
31-Aug-90	188			31-Aug-10	428		
30-Sep-90	189	63		30-Sep-10	429	143	
31-Oct-90	190			31-Oct-10	430		
30-Nov-90	191			30-Nov-10	431		
31-Dec-90	192	64	16	31-Dec-10	432	144	36
31-Jan-91	193			31-Jan-11	433		
28-Feb-91	194			28-Feb-11	434		
31-Mar-91	195	65		31-Mar-11	435	145	
30-Apr-91	196			30-Apr-11	436		
31-May-91	197			31-May-11	437		
30-Jun-91	198	66		30-Jun-11	438	146	
31-Jul-91	199			31-Jul-11	439		
31-Aug-91	200			31-Aug-11	440		
30-Sep-91	201	67		30-Sep-11	441	147	
31-Oct-91	202			31-Oct-11	442		
30-Nov-91	203			30-Nov-11	443		
31-Dec-91	204	68	17	31-Dec-11	444	148	37
31-Jan-92	205			31-Jan-12	445		
29-Feb-92	206			29-Feb-12	446		
31-Mar-92	207	69		31-Mar-12	447	149	
30-Apr-92	208			30-Apr-12	448		
31-May-92	209			31-May-12	449		
30-Jun-92	210	70		30-Jun-12	450	150	
31-Jul-92	211			31-Jul-12	451		
31-Aug-92	212			31-Aug-12	452		
30-Sep-92	213	71		30-Sep-12	453	151	
31-Oct-92	214			31-Oct-12	454		
30-Nov-92	215			30-Nov-12	455		
31-Dec-92	216	72	18	31-Dec-12	456	152	38
31-Jan-93	217			31-Jan-13	457		
28-Feb-93	218			28-Feb-13	458		

Reference date	Month (m)	Quarter (q)	Annual (a)	Reference date	Month (m)	Quarter (q)	Annual (a)
31-Mar-93	219	73		31-Mar-13	459	153	
30-Apr-93	220			30-Apr-13	460		
31-May-93	221			31-May-13	461		
30-Jun-93	222	74		30-Jun-13	462	154	
31-Jul-93	223			31-Jul-13	463		
31-Aug-93	224			31-Aug-13	464		
30-Sep-93	225	75		30-Sep-13	465	155	
31-Oct-93	226			31-Oct-13	466		
30-Nov-93	227			30-Nov-13	467		
31-Dec-93	228	76	19	31-Dec-13	468	156	39
31-Jan-94	229			31-Jan-14	469		
28-Feb-94	230			28-Feb-14	470		
31-Mar-94	231	77		31-Mar-14	471	157	
30-Apr-94	232			30-Apr-14	472		
31-May-94	233			31-May-14	473		
30-Jun-94	234	78		30-Jun-14	474	158	
31-Jul-94	235			31-Jul-14	475		
31-Aug-94	236			31-Aug-14	476		
30-Sep-94	237	79		30-Sep-14	477	159	
31-Oct-94	238			31-Oct-14	478		
30-Nov-94	239			30-Nov-14	479		
31-Dec-94	240	80	20	31-Dec-14	480	160	40