FDZ-METHODENREPORT

Methodological aspects of labour market data

01|2020 EN AKM effects for German labour market data

Lisa Bellmann, Ben Lochner, Stefan Seth, Stefanie Wolter

Bundesagentur für Arbeit

AKM effects for German labour market data

Lisa Bellmann (Institute for Employment Research)

Ben Lochner (Institute for Employment Research and Friedrich Alexander University Nuremberg)

Stefan Seth (Institute for Employment Research)

Stefanie Wolter (Institute for Employment Research)

Content

Tables

Appendix

Zusammenfassung

Dieser FDZ-Methodenreport beschreibt die Inhalte und Struktur der personen- und betriebsspezifischen Effekte und wie diese zu einigen der über das Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (IAB) verfügbaren Datensätze zugespielt werden können. Dieser FDZ-Methodenbericht ist eine Aktualisierung von Card et al. 2015.

Abstract

This FDZ Methodenreport describes content and structure of the person and establishment effects estimated and how this information may be combined with several of the data products available at the research Data Center (FDZ) of the German Federal Employment Agency at the Institute for Employment Research (IAB). This FDZ Methodenbericht updates Card et al. 2015.

Keywords

1   Introduction

In their 2013 paper, Card, Heining, and Kline (henceforce CHK) study the role of establishment-specific wage premia in generating recent increases in West German wage inequality. For five sub-intervals covering the period from 1985 to 2017, they estimate models with additive fixed effects for workers and establishments following the estimation strategy introduced in Abowd, Kramaz, and Margolis 1999 (hereafter AKM).

The analyses of Card et al. 2013 are based on administrative data for Germany, specifically the IAB Employment History File (BEH). This data set represents the universe of workers subject to social security contributions in Germany and contains information on employment and wages on a daily basis. The BEH is maintained by the Institute for Employment Research (IAB) and is one of the major sources of the data products available at the Research Data Center (FDZ) of the German Federal Employment Agency.

In the framework applied by Card et al. (2013), the estimates for person fixed effects may be interpreted as a combination of individual skills and other factors which are rewarded equally across employers. On the other hand, an establishment fixed effect represents the proportional pay premium (or discount) that is paid by specific establishments to all employees.

This paper describes the update and extension of the person and establishment fixed effects (AKM effects). For the original version refer to Card et al. 2015. The update covers the five periods 1985-1992, 1993 – 1999, 1998 – 2004, 2003 – 2010 and 2010 – 2017. The first period refers only to West Germany, while for the other periods estimates of the person and establishment fixed effects are available for both East and West. The FDZ has prepared several files containing the AKM effects, which may be linked to the Sample of Integrated Labor Market Biographies (SIAB) data, the Linked Employer-Employee Data (LIAB) of IAB and the ADIAB of the Linked-Employer-Employer Panel (LPP-ADIAB).

This FDZ Methodenreport describes these files and is structured as follows: Section 2 summarizes the changes from the original version of the effects. Section 3 provides a more detailed overview of the BEH and the sample used by Card et al. (2013). Their estimation strategy is briefly discussed in section 4. The linkage of the AKM effects to the SIAB data is described in section 5. Section 6 provides information on how to access these files. Finally, section 7 contains a description of variables.

2   Summary of changes

The update of the AKM is closely linked to the original version suggested by Card et al (2013). The following improvements and changes have been implemented:

- Extension of periods: the 5 periods now cover 1985-1992, 1993 – 1999, 1998 – 2004, 2003 – 2010 and 2010 – 2017.

- Extension to East Germany: periods two to five cover East and West Germany.
- Extension to women: the update includes males and females jointly. The original version included data for both sexes separately.

3 Data

As mentioned above, the BEH comprises information on employment subject to social security contributions. All information is available on a daily basis, which allows the construction of accurate work biographies. With regard to the AKM effects it is important that the data contain a worker ID as well as an establishment ID so that it is possible to track changing employers of a worker over time.

Workers can have several jobs at the same time with one or more employers. To select one observation per worker and year, we first average a worker's within-firm wages (if necessary), then we select the highest-paying employer (if necessary).

The update covers the five periods 1985-1992, 1993–1999, 1998–2004, 2003–2010, and 2010–2017. The first period refers to West Germany only; for periods two to five, we estimate the person and establishment fixed effects (AKM effects) for East and West jointly. We restrict our sample to full-time workers (person group 101) and apprentices (person group 102) aged between 20 and 60 and discard jobs with (real) daily wages below 10 Euros. Table 3-1 reports the number of estimated person and establishment fixed effects per period. Additional summary statistics can be found in the appendix.

Table 3-1 Number of estimated AKM effects for individuals and establishments

	1985-1992	1993-1999	1998-2004	2003-2010	2010-2017
Number of person effects	28,297,724	32,645,910	30,598,327	29,865,417	30,787,607
Number of establishment effects	1,898,388	2,543,452	2,537,182	2,476,096	2,103,298

4 Estimation

Before estimating the worker and establishment fixed effects, we impute wages above the social security contribution threshold. The imputation strategy follows Card et al. (2013), see the CHK paper's online appendix for details. We estimate tobit regressions by year, sex, education and age group. As controls we include: age, mean log wage in other years, fraction of censored wages in other years, number of full time employees at the current establishment and its square, dummy for 11 or more employees, mean years of schooling and fraction of university graduates at the current establishment, mean log wage of co-workers and fraction of co-workers with censored wages,

dummy for individuals observed only 1 year between 1985 and 2009, dummy for employees of 1-worker establishments, and an east dummy. For the estimation of worker and establishment fixed effects, Card et al. followed the frame-work introduced by Abowd et al in 1999.

$$y_{it} = \alpha_i + \psi_{J(i,t)} + x'_{it}\beta + r_{it}$$

In a given time interval the data set contains N^* person-year observations on N workers and J establishments. The function $J(i, t)$ gives the identity of the unique establishment that employs worker i in year t. It is assumed that the log daily real wage y_{it} of individual i in year t is the sum of a worker component α_i, an establishment component $\psi_{J(i,t)}$, an index of time-varying observable characteristics $x'_{it}\beta$, and an error component r_{it}. The person effect α_i is interpreted as a combination of skills and other factors that are rewarded equally across employers. Likewise, the index $x'_{it}\beta$ is interpreted as a combination of life cycle and aggregate factors that affect worker i's productivity at all jobs. An unrestricted set of year dummies as well as quadratic and cubic terms in age fully interacted with educational attainment is included in x_{it}. Finally, the establishment effect ψ_j is interpreted as a proportional pay premium (or discount) that is paid by establishment j to all employees (i.e., all those with $J(i, t) = j$. Such a premium could represent rent-sharing, an efficiency wage premium, or strategic wage posting behavior.

5   AKM Effects and Linkage

The data are stored in two different datasets, one containing establishment fixed effects and one containing person fixed effects. For IAB Establishment Panel and Establishment History Panel only establishment effects are available.

These files can be merged to the following datasets:

- Sample of Integrated Labor Market Biographies (Weakly anonymous version, SIAB 7517, Antoni et al. 2019)
- Linked Employer-Employee Data: Longitudinal Model (LIAB LM 7517, Schmidtlein et al. 2019a)
- Linked Employer-Employee Data: Cross-Sectional Model (LIAB QM2 9317, Schmidtlein et al. 2019b).
- LPP ADIAB (LPP-ADIAB 7517, Mackeben et al. 2019)
- IAB Establishment Panel (IAB-BP 9318)
- Establishment History Panel (BHP 7518, Ganzer et al. 2020).

The extension data sets are named as follows:

- **SIAB 7517**: SIAB_7517_v1_akm_pers.dta and SIAB_7517_v1_akm_estab.dta
- **LIAB LM 7517**: liab_lm_7517_v1_akm_pers.dta and liab_lm_7517_v1_akm_estab.dta
- **LIAB QM2 9317**: liab_qm2_9317_v1_akm_pers.dta and liab_qm2_9317_v1_akm_estab.dta
- **LPP ADIAB 7517:** lpp-adiab_7517_v1_akm_pers.dta and lpp-adiab_7517_v1_akm_estab.dta
- **IABBP 9318:** iabbp_9318_v1_akm_estab.dta
- **BHP 7518:** bhp_7518_m06_p50_v1_akm_estab.dta

The AKM person and establishment effects have been estimated only for regularly employed, full time workers, excluding for example, marginal workers. As consequence, the AKM effects may only be meaningfully interpreted for codes 101 and 102 of the *erwstat* variable **and** for the codes 0 of the teilzeit variable (`inlist(teilzeit, 0) & inlist(erwstat,101,102)`).

However, the merging procedure described below links AKM effects via persnr (person identifier) where the *erwstat* and *teilzeit* variables may have different values. It is up to the researcher whether to use the information for these observations or not.

The prepared files not only contain the point estimates of the person and establishment effects. Additionally, the individual and establishment fixed effects are ranked according to size and grouped into 20 quantiles indicating the position in the overall distribution of point estimates for both the person and the establishment effects are included.

Establishment and individual level file contain variable labels in German and English. To change the label language you can type `label language de/en` in Stata.

5.1 Exemplary Linkage to the Sample of Integrated Labor Market Biographies (SIAB 7517)

Exemplary to the merge with the SIAB data, the AKM effects can be linked to other datasets listed above following the same procedure. The AKM effects may be linked to the SIAB 7517 data (see Antoni et al. 2019) by using the *persnr* variable in order to link the person AKM effect respectively the or the *betnr* variable to link the establishment effects. To restrict the link to observations that overlap with the different periods in the AKM data, the researcher first needs to set up the SIAB data in annual structure and generate five AKM periods. An example how to generate annual data is provided in Eberle and Schmucker (2019).

```
use SIAB_7517_v1.dta, clear

merge m:1 persnr using SIAB_7517_v1_akm_pers_v1.dta
```

```
use SIAB_7517_v1.dta, clear

merge m:1 betnr using SIAB_7517_v1_akm_estab_v1.dta
```

Table 5-1 and Table 5-2 provide an overview on the number linked AKM effects in the SIAB 7517 data. The linkage rate for the unrestricted sample decreases over time, which is mainly driven by the increase of non-fulltime jobs.

Table 5-1 Number of linked AKM - Establishment effects in the SIAB 7517 data

	No restrictions		Restricted to: only BeH-spells, fulltime, age 20-60	
Interval	Total	Percentage	Total	Percentage

1985 – 1992	4,788,628	79.1	3,902,840	95.6
1993 – 1999	5,635,980	65.4	4,442,044	99.2
1998 – 2004	7,257,698	53.2	4,423,997	98.6
2003 – 2010	9,433,973	43.8	5,750,973	98.3
2010 – 2017	10,212,871	47.3	5,691,014	98.0

Table 5-2 Number of linked AKM-Person effects on the SIAB 7519 data

	No restrictions		Restricted to: only BeH-spells, fulltime, age 20-60	
Interval	Total	Percentage	Total	Percentage
1985 – 1992	4,234,136	69.7	3,726,392	100
1993 – 1999	5,606,592	65.1	4,201,169	100
1998 – 2004	7,463,446	54.7	4,449,821	100
2003 – 2010	10,234,750	47.6	5,233,233	100
2010 – 2017	9,391,414	43.5	5,225,135	100

6   Data access

The AKM effects will be provided to researchers working with the SIAB 7517, LIAB LM 9317, LIAB QM2 9317 or LPP ADIAB 7517 data upon request via e-mail by FDZ. An unexpired use agreement covering the respective data is required.

7   Description of Variables

7.1   Identifiers

7.1.1   Individual ID (persnr)

Variable label	Indiviudal ID
Variable name	persnr
Category	Identifier
Origin	BeH
Data type	Numerical
Hierarchy	None
Detailed description	The artificial individual ID indicates which observations belong to the same person. Artificial means that it is not possible to infer any of the person's characteristics or any original identifiers from this individual ID.

7.1.2   Establishment ID (betnr)

Variable label	Establishment ID
Variable name	betnr
Category	Identifier
Origin	BeH
Data type	Numerical
Hierarchy	None
Detailed description	Artificial establishment ID that is used for all administrative data products provided by FDZ.

7.2   Personal Variables

7.2.1   Person Effect 1985 – 1992 (peff_1985_1992)

Variable label	Person Effekt 1985-1992
Variable name	peff_1985-1992
Category	Personal variable
Origin	BeH
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the estimated person effects for both male and female workers in the period 1985 to 1991

7.2.2   Position in distribution of person effect 1985 – 1992 (p_peff_1985_1992)

Variable label	Position in distribution of person effect 1985-1992
Variable name	p_peff_1985-1992
Category	Personal variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the position of the person effect in the overall distributions of person effects for female and male workers in the period 1985 to 1992 in 20 quantiles.

7.2.3   Person Effect 1993 – 1999 (peff_1993_1999)

Variable label	Person Effekt 1993-1999
Variable name	peff_1993-1999
Category	Personal variable
Origin	BeH
Data type	Numerical
Hierarchy	None

Detailed description	This variable contains the estimated person effects for both male and female workers in the period 1993 to 1999.

7.2.4   Position in distribution of person effect 1993 – 1999 (p_peff_1993_1999)

Variable label	Position in distribution of person effect
Variable name	p_peff_1993-1999
Category	Personal variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the position of the person effect in the overall distributions of person effects for female and male workers in the period 1993 to 1999 in 20 quantiles.

7.2.5   Person Effect 1998 – 2004 (peff_1998_2004)

Variable label	Person Effekt 1998-2004
Variable name	peff_1998-2004
Category	Personal variable
Origin	BeH
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the estimated person effects for both male and female workers in the period 1998 to 2004.

7.2.6   Position in distribution of person effect 1998 – 2004 (p_peff_1998_2004)

Variable label	Position in distribution of person effect 1998-2004
Variable name	p_peff_1998-2004
Category	Personal variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the position of the person effect in the overall distributions of person effects for female and male workers in the period 1998 to 2004 in 20 quantiles.

7.2.7   Person Effect 2003 – 2010 (peff_2003_2010)

Variable label	Person Effekt 2003-2010
Variable name	peff_2003-2010
Category	Personal variable
Origin	BeH
Data type	Numerical
Hierarchy	None

Detailed description	This variable contains the estimated person effects for both male and female workers in the period 2003 to 2010.

7.2.8 Position in distribution of person effect 2003 – 2010 (p_peff_2003_2010)

Variable label	Position in distribution of person effect 2003-2010
Variable name	p_peff_2003-2010
Category	Personal variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the position of the person effect in the overall distributions of person effects for female and male workers in the period 2003 to 2010 in 20 quantiles.

7.2.9 Person Effect 2010 – 2017 (peff_2010_2017)

Variable label	Person Effekt 2010-2017
Variable name	peff_2010-2017
Category	Personal variable
Origin	BeH
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the estimated person effects for both male and female workers in the period 2010 to 2017.

7.2.10 Position in distribution of person effect 2010 – 2017 (p_peff_2010_2017)

Variable label	Position in distribution of person effect 2010-2017
Variable name	p_peff_2010-2017
Category	Personal variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the position of the person effect in the overall distributions of person effects for female and male workers in the period 2010 to 2017 in 20 quantiles.

7.3 Establishment Variables

7.3.1 Establishment Effect 1985 – 1992 (feff_1985-1992)

Variable label	Establishment Effect 1985-1992
Variable name	feff_1985-1992
Category	Establishment variable
Data type	Numerical

Hierarchy	None
Detailed description	This variable contains the estimated establishment effects for the period 1985 to 1991.

7.3.2 Position in distribution of establishment effects 1985 – 1992 (p_feff_1985-1992)

Variable label	Position in distribution of establishment effects 1985-1992
Variable name	p_feff_1985-1992
Category	Establishment variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the position of the establishment effect in the overall distribution of establishment effects for female and male workers in the period 1985 to 1991 in 20 quantiles.

7.3.3 Establishment Effect 1993 – 1999 (feff_1993_1999)

Variable label	Establishment Effect 1993-1999
Variable name	feff_1993-1999
Category	Establishment variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the estimated establishment effects for the period 1993 to 1999.

7.3.4 Position in distribution of establishment effects 1993 – 1999 (p_feff_1993_1999)

Variable label	Position in distribution of establishment effects 1993-1999
Variable name	p_feff_1993-1999
Category	Establishment variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the position of the establishment effect in the overall distribution of establishment effects for female and male workers in the period 1993 to 1999 in 20 quantiles.

7.3.5 Establishment Effect 1998 – 2004 (feff_1998_2004)

Variable label	Establishment Effect 1998-2004
Variable name	feff_1998-2004
Category	Establishment variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the estimated establishment effects for the period 1998 to 2004.

7.3.6 Position in distribution of establishment effects 1998 – 2004 (p_feff_1998_2004)

Variable label	Position in distribution of establishment effects 1998_2004
Variable name	p_feff_1998-2004
Category	Establishment variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the position of the establishment effect in the overall distribution of establishment effects for female and male workers in the period 1998 to 2004 in 20 quantiles.

7.3.7 Establishment Effect 2003 – 2010 (feff_2003_2010)

Variable label	Establishment Effect 2003-2010
Variable name	feff_2003-2010
Category	Establishment variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the estimated establishment effects for the period 2003 to 2010.

7.3.8 Position in distribution of establishment effects 2003 – 2010 (p_feff_2003_2010)

Variable label	Position in distribution of establishment effects 2003-2010
Variable name	p_feff_2003-2010
Category	Establishment variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the position of the establishment effect in the overall distribution of establishment effects for female and male workers in the period 2003 to 2010 in 20 quantiles.

7.3.9 Establishment Effect 2010 – 2017 (feff_2010_2017)

Variable label	Establishment Effect 2010-2017
Variable name	feff_2010-2017
Category	Establishment variable
Data type	Numerical
Hierarchy	None
Detailed description	This variable contains the estimated establishment effects for the period 2010 to 2017.

7.3.10 Position in distribution of establishment effects 2010 – 2017 (feff_2010-2017)

Variable label	Position in distribution of establishment effects 2010-2017
Variable name	p_feff_2010-2017
Category	Establishment variable
Data type	Numerical

Hierarchy	None
Detailed description	This variable contains the position of the establishment effect in the overall distribution of establishment effects for female and male workers in the period 2010 to 2017 in 20 quantiles.

Literature

Abowd, John; Kramarz, Francis; Margolis, David (1999), 'High Wage Workers and High Wage Firms, In: Econometrica, Vol. 67, No. 2, pp. 251–333.

Antoni, Manfred; Schmucker, Alexandra; Seth, Stefan; Vom Berge, Philipp (2019): Sample of integrated labour market biographies (SIAB) 1975-2017. (FDZ-Datenreport, 02/2019 (en)), Nürnberg, 72 S.

Card, David; Heining, Jörg; Kline, Patrick (2013): Workplace heterogeneity and the rise of West German wage inequality. In: The Quarterly Journal of Economics, Vol. 128, No. 3, pp. 967-1015.

Card, David; Heining, Jörg; Kline, Patrick (2015): CHK effects. (FDZ-Methodenreport, 06/2015 (en)), Nürnberg, 35 S.

Eberle, Johanna; Schmucker, Alexandra (2019): Creating cross-sectional data and biographical variables with the Sample of Integrated Labour Market Biographies 1975-2017 * Programming examples for Stata. (FDZ-Methodenreport, 04/2019 (en)), Nürnberg, 16 S.

Ganzer, Andreas; Schmidtlein, Lisa; Stegmaier, Jens; Wolter, Stefanie (2020): Establishment History Panel 1975-2018. (FDZ-Datenreport, 01/2020 (en)), Nürnberg, 112 S.

Mackeben, Jan; Ruf, Kevin; Haepp, Tobias; Wolter, Stefanie; Grunau, Philipp (2019): Linked Personnel Panel survey data linked with administrative data of the IAB (LPP-ADIAB) 1975 - 2017. (FDZ-Datenreport, 07/2019 (en)), Nürnberg, 79 S.

Schmidtlein, Lisa; Seth, Stefan; Umkehrer, Matthias (2019a): Linked Employer-Employee Data from the IAB: LIAB Longitudinal Model (LIAB LM) 1975-2017. (FDZ-Datenreport, 05/2019 (en)), Nürnberg, 63 S.

Schmidtlein, Lisa; Seth, Stefan; Umkehrer, Matthias (2019b): Linked Employer-Employee Data from the IAB: LIAB Cross-Sectional Model 2 (LIAB QM2) 1993-2017. (FDZ-Datenreport, 06/2019 (en)), Nürnberg, 77 S.

Appendix

A1 Summary Statistics Person Effects

1985-1992

5%-Percentile Position	N	mean	sd	p25	p50	p75	p99
1	1,414,893	3.34386	0.28859	3.22905	3.43220	3.55493	3.63905
2	1,414,902	3.75022	0.05357	3.70689	3.75675	3.79708	3.82978
3	1,414,874	3.88494	0.02892	3.86051	3.88664	3.91024	3.93095
4	1,414,931	3.96929	0.02068	3.95165	3.97009	3.98734	4.00292
5	1,414,906	4.03291	0.01637	4.01886	4.03341	4.04721	4.05978
6	1,414,849	4.08454	0.01362	4.07284	4.08485	4.09638	4.10708
7	1,414,981	4.12841	0.01180	4.11826	4.12864	4.13868	4.14805
8	1,414,813	4.16705	0.01058	4.15793	4.16718	4.17624	4.18476
9	1,414,986	4.20231	0.00982	4.19382	4.20240	4.21083	4.21881
10	1,414,827	4.23545	0.00937	4.22735	4.23550	4.24356	4.25129
11	1,414,852	4.26761	0.00923	4.25960	4.26761	4.27560	4.28328
12	1,414,920	4.29977	0.00938	4.29164	4.29975	4.30787	4.31579
13	1,414,894	4.33300	0.00986	4.32446	4.33289	4.34152	4.34991
14	1,414,806	4.36858	0.01077	4.35922	4.36840	4.37786	4.38721
15	1,414,958	4.40855	0.01242	4.39771	4.40825	4.41924	4.43021
16	1,414,874	4.45605	0.01523	4.44276	4.45552	4.46907	4.48294
17	1,414,838	4.51672	0.02026	4.49900	4.51572	4.53394	4.55306
18	1,414,880	4.60242	0.03024	4.57592	4.60038	4.62792	4.65772
19	1,414,858	4.73866	0.05009	4.69457	4.73507	4.78103	4.83040
20	1,414,882	5.00667	0.16493	4.89124	4.96457	5.07509	5.53964
Total	28,297,724	4.23985	0.36039	4.06029	4.25160	4.43068	5.10801

1993-1999

5%-Percentile Position	N	mean	sd	p25	p50	p75	p99
1	1,632,298	3.37625	0.29992	3.25578	3.47058	3.59673	3.68124
2	1,632,307	3.78979	0.05175	3.74820	3.79667	3.83507	3.86561
3	1,632,307	3.91639	0.02660	3.89393	3.91798	3.93966	3.95869
4	1,632,348	3.99398	0.01904	3.97775	3.99470	4.01059	4.02499
5	1,632,223	4.05278	0.01520	4.03977	4.05320	4.06603	4.07782
6	1,632,450	4.10116	0.01286	4.09010	4.10144	4.11235	4.12249
7	1,632,191	4.14297	0.01139	4.13316	4.14313	4.15286	4.16200
8	1,632,249	4.18075	0.01048	4.17170	4.18086	4.18986	4.19835
9	1,632,445	4.21602	0.00994	4.20742	4.21607	4.22464	4.23281
10	1,632,141	4.25000	0.00971	4.24160	4.25002	4.25841	4.26645
11	1,632,385	4.28365	0.00977	4.27519	4.28362	4.29210	4.30030
12	1,632,208	4.31805	0.01014	4.30925	4.31799	4.32680	4.33542
13	1,632,313	4.35444	0.01094	4.34494	4.35429	4.36387	4.37330
14	1,632,381	4.39452	0.01230	4.38382	4.39425	4.40509	4.41591

15	1,632,284	4.44085	0.01457	4.42815	4.44045	4.45335	4.46638
16	1,632,250	4.49725	0.01824	4.48133	4.49657	4.51287	4.52948
17	1,632,263	4.56951	0.02377	4.54872	4.56857	4.58981	4.61170
18	1,632,317	4.66569	0.03246	4.63729	4.66401	4.69321	4.72419
19	1,632,262	4.80722	0.05149	4.76192	4.80332	4.85059	4.90192
20	1,632,288	5.06204	0.14593	4.96397	5.03239	5.12343	5.51838
Total	32,645,910	4.27066	0.36898	4.07831	4.26679	4.46693	5.14871

1998-2004

5%-Percentile Position	N	mean	sd	p25	p50	p75	p99
1	1,529,920	3.41033	0.30796	3.27552	3.50692	3.64131	3.73237
2	1,529,922	3.85319	0.05807	3.80621	3.86081	3.90410	3.93848
3	1,529,928	3.99470	0.02922	3.97016	3.99663	4.02027	4.04094
4	1,529,965	4.07885	0.02042	4.06141	4.07967	4.09668	4.11203
5	1,529,957	4.14148	0.01607	4.12771	4.14198	4.15548	4.16791
6	1,529,905	4.19245	0.01353	4.18083	4.19273	4.20423	4.21492
7	1,529,874	4.23654	0.01201	4.22619	4.23673	4.24697	4.25659
8	1,529,945	4.27640	0.01109	4.26682	4.27649	4.28602	4.29504
9	1,529,856	4.31384	0.01059	4.30469	4.31387	4.32303	4.33172
10	1,529,901	4.35009	0.01039	4.34109	4.35011	4.35910	4.36770
11	1,529,937	4.38620	0.01052	4.37708	4.38615	4.39529	4.40417
12	1,529,930	4.42351	0.01106	4.41391	4.42341	4.43306	4.44251
13	1,529,878	4.46345	0.01207	4.45296	4.46328	4.47385	4.48430
14	1,529,915	4.50791	0.01370	4.49599	4.50761	4.51968	4.53173
15	1,529,930	4.55967	0.01636	4.54538	4.55920	4.57368	4.58836
16	1,529,912	4.62301	0.02042	4.60519	4.62231	4.64048	4.65905
17	1,529,909	4.70326	0.02617	4.68040	4.70228	4.72563	4.74957
18	1,529,924	4.80810	0.03506	4.77745	4.80645	4.83791	4.87090
19	1,529,919	4.95553	0.05152	4.91029	4.95252	4.99918	5.04887
20	1,529,900	5.21825	0.15890	5.10970	5.18068	5.28343	5.73657
Total	30,598,327	4.37484	0.39716	4.16841	4.36806	4.58898	5.31358

2003-2010

5%-Percentile Position	N	mean	sd	p25	p50	p75	p99
1	1,493,274	3.72623	0.31267	3.57841	3.82278	3.96492	4.06002
2	1,493,306	4.18572	0.06042	4.13682	4.19355	4.23874	4.27448
3	1,493,235	4.33273	0.03027	4.30730	4.33471	4.35920	4.38057
4	1,493,276	4.41998	0.02128	4.40179	4.42082	4.43856	4.45462
5	1,493,296	4.48556	0.01691	4.47105	4.48604	4.50030	4.51336
6	1,493,332	4.53927	0.01430	4.52697	4.53954	4.55172	4.56303
7	1,493,208	4.58600	0.01279	4.57498	4.58615	4.59710	4.60742
8	1,493,305	4.62875	0.01196	4.61844	4.62884	4.63912	4.64889
9	1,493,320	4.66936	0.01153	4.65940	4.66941	4.67936	4.68886
10	1,493,194	4.70901	0.01142	4.69912	4.70900	4.71890	4.72841
11	1,493,325	4.74883	0.01160	4.73877	4.74880	4.75888	4.76861

12	1,493,187	4.78986	0.01215	4.77932	4.78975	4.80036	4.81069
13	1,493,314	4.83351	0.01313	4.82208	4.83332	4.84483	4.85614
14	1,493,226	4.88163	0.01477	4.86877	4.88132	4.89435	4.90731
15	1,493,304	4.93721	0.01748	4.92198	4.93675	4.95219	4.96779
16	1,493,246	5.00430	0.02152	4.98552	5.00359	5.02274	5.04216
17	1,493,273	5.08802	0.02705	5.06438	5.08707	5.11115	5.13578
18	1,493,291	5.19604	0.03621	5.16436	5.19422	5.22673	5.26121
19	1,493,238	5.35136	0.05538	5.30273	5.34763	5.39803	5.45263
20	1,493,267	5.63505	0.16817	5.52023	5.59742	5.70370	6.19869
Total	29,865,417	4.73742	0.42029	4.51389	4.72881	4.96846	5.73422

2010-2017

5%-Percentile Position	N	mean	sd	p25	p50	p75	p99
1	1,539,382	3.17946	0.30760	3.04316	3.27904	3.41119	3.49497
2	1,539,397	3.59803	0.04891	3.55882	3.60448	3.64079	3.66986
3	1,539,366	3.71914	0.02596	3.69721	3.72056	3.74182	3.76072
4	1,539,417	3.79668	0.01960	3.77991	3.79729	3.81375	3.82892
5	1,539,346	3.85894	0.01660	3.84465	3.85928	3.87338	3.88651
6	1,539,392	3.91311	0.01480	3.90033	3.91332	3.92596	3.93785
7	1,539,413	3.96236	0.01373	3.95051	3.96246	3.97427	3.98546
8	1,539,384	4.00890	0.01321	3.99745	4.00897	4.02035	4.03125
9	1,539,387	4.05428	0.01302	4.04301	4.05428	4.06556	4.07637
10	1,539,426	4.09946	0.01310	4.08812	4.09944	4.11079	4.12177
11	1,539,351	4.14543	0.01347	4.13376	4.14537	4.15708	4.16843
12	1,539,369	4.19326	0.01423	4.18092	4.19310	4.20553	4.21773
13	1,539,355	4.24477	0.01560	4.23121	4.24453	4.25821	4.27173
14	1,539,360	4.30235	0.01781	4.28684	4.30194	4.31766	4.33334
15	1,539,365	4.36941	0.02108	4.35102	4.36886	4.38752	4.40628
16	1,539,431	4.45000	0.02562	4.42764	4.44928	4.47201	4.49484
17	1,539,334	4.54778	0.03094	4.52078	4.54699	4.57437	4.60181
18	1,539,379	4.66578	0.03756	4.63303	4.66457	4.69790	4.73194
19	1,539,387	4.81494	0.04978	4.77143	4.81243	4.85709	4.90465
20	1,539,366	5.07051	0.15647	4.96374	5.03289	5.13226	5.60287
Total	30,787,607	4.14973	0.43482	3.88705	4.12223	4.40707	5.16146

A2 Summary Statistics Establishment Effects

1985-1992

5%-Percentile Position	N	mean	sd	p25	p50	p75	p99
1	94,920	-1.10804	0.28873	-1.25219	-1.02542	-0.88633	-0.78898
2	94,920	-0.64534	0.07212	-0.70522	-0.63849	-0.58199	-0.53470
3	94,919	-0.45380	0.04235	-0.48952	-0.45117	-0.41657	-0.38695
4	94,920	-0.33274	0.02896	-0.35743	-0.33138	-0.30756	-0.28581
5	94,918	-0.24548	0.02178	-0.26409	-0.24460	-0.22652	-0.21003
6	94,920	-0.17829	0.01732	-0.19310	-0.17775	-0.16322	-0.14981
7	94,919	-0.12372	0.01433	-0.13596	-0.12327	-0.11127	-0.10009
8	94,920	-0.07794	0.01223	-0.08852	-0.07761	-0.06729	-0.05768
9	94,919	-0.03814	0.01086	-0.04753	-0.03797	-0.02872	-0.02006
10	94,919	-0.00251	0.00982	-0.01098	-0.00243	0.00601	0.01396
11	94,920	0.03035	0.00918	0.02243	0.03043	0.03831	0.04578
12	94,920	0.06132	0.00874	0.05374	0.06141	0.06886	0.07608
13	94,919	0.09123	0.00855	0.08381	0.09129	0.09865	0.10569
14	94,920	0.12063	0.00851	0.11327	0.12058	0.12799	0.13515
15	94,918	0.15033	0.00872	0.14274	0.15025	0.15787	0.16527
16	94,920	0.18152	0.00935	0.17340	0.18138	0.18956	0.19768
17	94,919	0.21615	0.01084	0.20670	0.21581	0.22546	0.23516
18	94,921	0.25866	0.01402	0.24638	0.25801	0.27050	0.28385
19	94,918	0.32006	0.02293	0.29998	0.31800	0.33892	0.36375
20	94,919	0.52157	0.20236	0.39983	0.45306	0.56141	1.35893
Total	1,898,388	-0.06271	0.36764	-0.20938	0.01428	0.16560	0.60161

1993-1999

5%-Percentile Position	N	mean	sd	p25	p50	p75	p99
1	127,173	-1.03543	0.29323	-1.16876	-0.94575	-0.81501	-0.72529
2	127,174	-0.59678	0.06396	-0.64962	-0.58984	-0.54052	-0.50017
3	127,171	-0.43199	0.03569	-0.46211	-0.42969	-0.40081	-0.37548
4	127,173	-0.32930	0.02478	-0.35037	-0.32822	-0.30770	-0.28913
5	127,172	-0.25349	0.01948	-0.27024	-0.25289	-0.23650	-0.22161
6	127,173	-0.19202	0.01632	-0.20611	-0.19168	-0.17780	-0.16497
7	127,173	-0.13916	0.01434	-0.15156	-0.13893	-0.12668	-0.11521
8	127,172	-0.09220	0.01284	-0.10329	-0.09203	-0.08104	-0.07066
9	127,173	-0.04949	0.01185	-0.05975	-0.04935	-0.03916	-0.02966
10	127,172	-0.00978	0.01118	-0.01943	-0.00973	-0.00009	0.00908
11	127,173	0.02794	0.01062	0.01875	0.02802	0.03717	0.04584
12	127,173	0.06400	0.01020	0.05519	0.06405	0.07283	0.08122
13	127,172	0.09879	0.00991	0.09022	0.09884	0.10739	0.11552
14	127,173	0.13283	0.00978	0.12436	0.13285	0.14129	0.14942
15	127,172	0.16680	0.00990	0.15819	0.16675	0.17534	0.18376
16	127,173	0.20197	0.01045	0.19288	0.20187	0.21101	0.21991
17	127,173	0.24001	0.01167	0.22989	0.23971	0.25002	0.26042

5%-Percentile Position	N	mean	sd	p25	p50	p75	p99
18	127,172	0.28467	0.01439	0.27210	0.28409	0.29690	0.31031
19	127,173	0.34639	0.02280	0.32648	0.34416	0.36518	0.39002
20	127,172	0.55020	0.20909	0.42611	0.47903	0.58780	1.43145
Total	2,543,452	-0.05080	0.36083	-0.22096	0.00946	0.18410	0.62935

1998-2004

5%-Percentile Position	N	mean	sd	p25	p50	p75	p99
1	126,860	-1.20610	0.31546	-1.36045	-1.11255	-0.96498	-0.86357
2	126,859	-0.72311	0.06982	-0.78065	-0.71579	-0.66186	-0.61750
3	126,859	-0.54367	0.03834	-0.57591	-0.54111	-0.51010	-0.48315
4	126,859	-0.43453	0.02606	-0.45674	-0.43340	-0.41171	-0.39239
5	126,860	-0.35521	0.02024	-0.37251	-0.35464	-0.33757	-0.32198
6	126,858	-0.29146	0.01678	-0.30586	-0.29112	-0.27687	-0.26358
7	126,859	-0.23677	0.01484	-0.24956	-0.23645	-0.22383	-0.21217
8	126,860	-0.18809	0.01341	-0.19965	-0.18789	-0.17651	-0.16557
9	126,858	-0.14349	0.01235	-0.15412	-0.14339	-0.13277	-0.12274
10	126,859	-0.10214	0.01157	-0.11214	-0.10209	-0.09205	-0.08270
11	126,860	-0.06304	0.01110	-0.07265	-0.06298	-0.05335	-0.04435
12	126,859	-0.02533	0.01073	-0.03464	-0.02527	-0.01604	-0.00720
13	126,859	0.01147	0.01047	0.00243	0.01155	0.02053	0.02910
14	126,859	0.04733	0.01032	0.03839	0.04735	0.05623	0.06490
15	126,859	0.08328	0.01042	0.07421	0.08332	0.09224	0.10103
16	126,859	0.12003	0.01090	0.11058	0.11993	0.12942	0.13874
17	126,859	0.15965	0.01214	0.14909	0.15938	0.17013	0.18078
18	126,859	0.20615	0.01504	0.19308	0.20549	0.21888	0.23312
19	126,859	0.27116	0.02408	0.25010	0.26884	0.29093	0.31717
20	126,859	0.49150	0.22724	0.35551	0.41260	0.53235	1.44823
Total	2,537,182	-0.14612	0.38348	-0.32135	-0.08231	0.10138	0.57736

2003-2010

5%-Percentile Position	N	mean	sd	p25	p50	p75	p99
1	123,805	-1.73559	0.31647	-1.90762	-1.65721	-1.48756	-1.36649
2	123,805	-1.19431	0.08545	-1.26451	-1.18538	-1.11938	-1.06493
3	123,805	-0.97333	0.04760	-1.01330	-0.97015	-0.93165	-0.89834
4	123,807	-0.83808	0.03206	-0.86529	-0.83638	-0.81006	-0.78657
5	123,802	-0.74238	0.02395	-0.76295	-0.74153	-0.72148	-0.70334
6	123,806	-0.66818	0.01923	-0.68463	-0.66766	-0.65139	-0.63644
7	123,804	-0.60660	0.01651	-0.62075	-0.60624	-0.59224	-0.57913
8	123,805	-0.55275	0.01470	-0.56539	-0.55253	-0.53996	-0.52816
9	123,806	-0.50401	0.01351	-0.51564	-0.50391	-0.49229	-0.48135
10	123,803	-0.45886	0.01261	-0.46978	-0.45877	-0.44791	-0.43760
11	123,805	-0.41631	0.01194	-0.42664	-0.41628	-0.40589	-0.39624
12	123,805	-0.37590	0.01147	-0.38579	-0.37587	-0.36596	-0.35649
13	123,806	-0.33670	0.01120	-0.34640	-0.33670	-0.32702	-0.31768

5%-Percentile Position	N	mean	sd	p25	p50	p75	p99
14	123,805	-0.29774	0.01132	-0.30756	-0.29779	-0.28792	-0.27851
15	123,803	-0.25831	0.01148	-0.26822	-0.25838	-0.24841	-0.23866
16	123,805	-0.21735	0.01222	-0.22799	-0.21751	-0.20682	-0.19636
17	123,805	-0.17268	0.01374	-0.18465	-0.17303	-0.16088	-0.14871
18	123,806	-0.11938	0.01748	-0.13467	-0.12019	-0.10451	-0.08825
19	123,804	-0.04436	0.02758	-0.06845	-0.04682	-0.02154	0.00798
20	123,804	0.19644	0.23995	0.05078	0.11294	0.24304	1.18033
Total	2,476,096	-0.51582	0.43532	-0.70263	-0.43718	-0.23825	0.29247

2010-2017

5%-Percentile Position	N	mean	sd	p25	p50	p75	p99
1	105,165	-0.95103	0.33636	-1.14400	-0.85880	-0.68280	-0.56744
2	105,165	-0.41474	0.07462	-0.47593	-0.40563	-0.34920	-0.30388
3	105,165	-0.22988	0.03803	-0.26162	-0.22697	-0.19662	-0.17039
4	105,165	-0.12450	0.02433	-0.14513	-0.12308	-0.10328	-0.08555
5	105,165	-0.05236	0.01799	-0.06774	-0.05175	-0.03663	-0.02307
6	105,165	0.00333	0.01445	-0.00905	0.00365	0.01594	0.02724
7	105,165	0.04988	0.01254	0.03911	0.05010	0.06079	0.07071
8	105,165	0.09078	0.01120	0.08113	0.09087	0.10051	0.10963
9	105,165	0.12813	0.01039	0.11916	0.12819	0.13717	0.14562
10	105,164	0.16304	0.00979	0.15456	0.16308	0.17155	0.17954
11	105,165	0.19631	0.00943	0.18815	0.19634	0.20449	0.21229
12	105,166	0.22876	0.00932	0.22067	0.22876	0.23682	0.24454
13	105,165	0.26100	0.00933	0.25290	0.26099	0.26907	0.27688
14	105,164	0.29362	0.00953	0.28536	0.29354	0.30189	0.30987
15	105,165	0.32731	0.00996	0.31866	0.32723	0.33596	0.34434
16	105,165	0.36302	0.01074	0.35373	0.36284	0.37224	0.38161
17	105,165	0.40262	0.01223	0.39199	0.40230	0.41313	0.42401
18	105,165	0.45064	0.01596	0.43670	0.44989	0.46420	0.47918
19	105,165	0.51972	0.02550	0.49736	0.51745	0.54075	0.56808
20	105,164	0.74158	0.22557	0.60667	0.66394	0.78171	1.70096
Total	2,103,298	0.12236	0.37050	-0.02249	0.17988	0.34469	0.82687

Imprint

Corresponding author
Stefanie Wolter
Phone: +49 911 179-1615
Email stefanie.wolter@iab.de