Data Observer

Steffi Dierks*, Alexander Schiersch and Jan Stede

Industry Conversion Tables for German Firm-Level Data

<https://doi.org/10.1515/jbnst-2019-0041>

1 Introduction

The relevance of microdata for economic analysis and policy evaluation has grown rapidly in the last decade. One reason for this, is that ever more data is collected and existing microdata are made available to researchers. The establishment of the Research Data Center (FDZ) was a milestone in this process in Germany. Located in almost every federal state, they grant easy access to the enormous treasure of official microdata (Zühlke et al. 2004). The creation of the *AFiD-Panels*, the German acronym for “official firm-level data in Germany,” also contributes to the accessibility of firm-level data by combining different official surveys and censuses. For more information on the *AFiD-Panels*, we refer to Wagner (2000, 2008, 2010), Fritsch et al. (2004), Konold (2007), Malchin and Voshage (2009), and Vogel (2009).

There are several reasons why German firm-level data is of high quality compared to other sources. First, firms are legally obliged to deliver the requested data. Moreover, firms have almost no option to refuse participation once they are selected into a survey. Both reasons lead to extremely low non-response rates, e. g. less than two percent in the case of the cost structure survey of manufacturing firms (Statistisches Bundesamt 2016). Third, the data are collected and processed by the staff of the Statistical Offices. This ensures very high data quality, since these institutions have comparatively larger resources than private firms and its staff are well qualified. Finally, the data are usually collected such that it is representative at the industry level and, often, also at the state level.

***Corresponding author: Steffi Dierks**, Statistisches Amt für Hamburg und Schleswig-Holstein, Fröbelstraße 15-17, Kiel 24113, Germany, E-mail: steffi.dierks@statistik-nord.de

Alexander Schiersch: E-mail: aschiersch@diw.de, **Jan Stede**: E-mail: jstede@diw.de, DIW Berlin, German Institute for Economic Research, Mohrenstraße 58, Berlin 10117, Germany

In applied research using firm-level data, it is often imperative to have a panel with an industry classification that is consistent over time. Monetary variables, for example, need to be deflated in order to allow comparisons over time. The estimation of production functions should also be conducted on a disaggregated industry level (if sufficient observations are available), usually defined by the two-digit or three-digit industry code. This follows from the assumption typically imposed in production function approaches that all firms operate within the same production possibility set, using the same technology. This assumption is much more credible when the estimation is done at the disaggregated industry level. Finally, dummies are often used to capture structural differences between industries.

The use of such firm-level data is hampered, however, by institutional changes, including the change of industry classification adjusting to changes in economic structures. In particular, the change from ISIC Rev. 3.1 to ISIC Rev. 4, implemented in Germany as the change from WZ 2003 to WZ 2008 (hereinafter: WZ03 and WZ08, respectively),¹ led to a disruption in the time series. Unlike previous changes, it was not just a relabeling of industries: the introduction of some new sections (e. g. the new created section *information and communication*) makes it difficult to compare data between WZ03 and WZ08 (Greulich 2009).² In addition, entire sectors were dissolved and the companies that were part of these sectors were assigned to newly established or already existing industries.³ This is not only true for two-digit industries within one-digit industries, such as in manufacturing, but firms also moved between one-digit industries. For example, industry D22 (*Publishing, printing and reproduction of recorded media*) is part of manufacturing according to WZ03, but large parts of it moved to J58 (*Publishing activities*) in WZ08 and, therefore, is no longer classified as manufacturing in the new system. Disregarding such inconsistencies in the sectoral affiliation of firms may lead to a situation where, for example, two different sectoral deflators are used for different periods of time for a single firm that has not changed its main economic activity.

¹ The International Standard Industrial Classification of All Economic Activities (ISIC) is a widely used United Nations industry classification system. The European system, the Statistical Classification of Economic Activities in the European Community (NACE), is based on the ISIC. NACE Rev. 1.1 (building on ISIC Rev. 3) and NACE Rev. 2 (building on ISIC Rev. 4) are implemented as the German Classification of Economic Activities WZ03 and WZ08, respectively.

² The list of WZ 2008 sectors is available at Statistisches Bundesamt (2008a).

³ The term “sector” in this study does not refer to the sector as used in national accounts, such as state, households etc. We use the terms ‘sector’ and ‘industry’ synonymously.

The fundamental problem of the industry reclassification is that there is no one-to-one mapping between old and new industries: firms from an old industry moved to different economic sectors after 2008 and new industries combine firms from different former industries. Conversion tables providing one-to-one mappings that can be universally applied are one solution to this issue. However, the official conversion tables published by the Federal Statistical Office of Germany cannot be used in applied research. These tables simply contain all possible combinations between old and new sectors, without specifying what fraction of a certain sector in WZ03 was assigned to which sector in WZ08.

In this paper, we provide ready-to-use conversion tables for German firm-level data.⁴ We apply a strategy to rematch firms between industry classification systems based on a propensity score. Specifically, we estimate the “most likely counterpart” of each economic sector by comparing the relative share of each industry combination/transition path in the overall number of switchings per industry observed in the German business register. We then select the industry combination/transition path with the highest share as the relevant one for the one-to-one-mapping. As a robustness check, we compare the estimated one-to-one mappings with the universe of industry combinations/transition paths published by the Federal Statistical Office of Germany.

We provide conversion tables for the five-digit, four-digit and three-digit industry code level. Four-digit sectoral codes are also included in the AFiD-Panels, which means our conversion tables can be directly applied to this official data as well as to any other German firm-level data with three-digit or more detailed industry codes. All conversion tables are provided for two directions, both to convert WZ03 codes to WZ08, and vice versa.

The remainder of this paper is organized as follows: Section 2 describes the data and method. Section 3 shows the results and Section 4 summarizes.

2 Data and methodology

We estimate the industry conversion tables using the *AFiD-Panel Business Register*. This dataset contains all economically active units – both single-plant firms and multi-plant companies – that provide a contribution to the German gross domestic product, have their registered office in Germany, and focus on one of the sections B (*Mining and quarrying*) to N (*Administrative and*

⁴ The tables are available for download from <https://bit.ly/2MPHBL8> (DOI:dx.doi.org/10.25652/diw_data_S0018.1).

support service activities), as well as P (*Education*) to S (*Other service activities*), of the statistical classification of economic activities in the European Community NACE Rev. 2 (Statistische Ämter des Bundes und der Länder 2015). The coverage of economic units is optional for section A (*Agriculture, forestry and fishing*), as well as O (*Public administration, defense and social insurance*). Sections T (*Activities of households as employers; undifferentiated goods- and service-producing activities of households for own use*) and U (*Activities of extraterritorial organizations and bodies*) are excluded.

The Business Register is a statistical register that is created and maintained using existing data from different sources. In Germany, these are mainly the files of the tax administration, the monthly results from the employment statistics of the Federal Employment Agency on companies and employees, as well as the annual files of the craft chambers and the chambers of industry and commerce. The most relevant items in the panel are the unique company and establishment identification numbers and the sector of economic activity according to the German industry classification WZ03 or WZ08.⁵

For calculating the conversion tables, the WZ03 and the WZ08 codes for each firm are required. Unfortunately, the data do not contain both codes for most firms within the year of the introduction of WZ08. Thus, for the creation of the industry conversion tables, we assume that firms do not change their main economic activity within two subsequent years. In other words, we assume that a company that produces cars at time t , for example, continues to be primarily a car manufacturer in $t + 1$.

For a conversion table to work, one unique WZ08 code for every single WZ03 code is needed. However, the change in the classification system was so significant that there is no one-to-one mapping between industries. Rather, firms from the same industry in 2007 moved to distinct industries after 2008 and industries in 2008 combine firms from different industries of 2007. Consequently, there are many different WZ08 codes for each WZ03 code and vice versa, even at the detailed 5-digit level.

We use the relative majority to overcome this problem, defining each pair in the conversion table by a WZ03 code and the respective WZ08 code that most of

⁵ Other variables include the name and address of the unit, an eight-digit regional code, its legal form (for legal entities), the number of employees subject to social insurance (for companies), its turnover, its membership numbers in chambers, and its tax number. Researchers can assess the Business Register – similar to other datasets from the AFID-Panels – either by remote access or (a reduced version) at a workplace in the research data center of the Statistical Offices of the Länder. Although the Business Register can be used for analyses, its main purpose is to serve as a basis from which the relevant units for the specific statistics are drawn.

the firms were assigned to. In order to construct the relative majority, we focus on the years 2007 and 2008 and count those firms with different sectoral codes in 2008 (the first year in which WZ08 codes are used in the AfID-Panel Business Register) that were in the same sector in 2007 (the latest year that WZ03 codes were provided). Dividing these numbers by the overall number of observations in a specific WZ03 sector provides shares. Finally, the WZ08 code with the highest share per ISIC 3.1 code is selected for the conversion table. The same approach is applied, just with the opposite direction, when constructing of the tables for the conversion from WZ08 to WZ03.

Example. Table 1 provides a stylized example of our methodology. Different companies that were all in the same sector in 2007 (sector 01132) are assigned to different sectors in 2008. Most companies – 75 % in the example – are assigned to sector 01210. Consequently, in our conversion table all companies of WZ03 sector 01132 are assumed to be assigned to the WZ08 sector 01210.

Table 1: Stylized example of the methodology used to construct the conversion tables.

Company ID	WZ03 (in 2007)	WZ08 (in 2008)	N ^{WZ03}	N ^{WZ08}	Max(share)
0000001	01132	01210	8	6	75 %
1230012	01132	01210	8	6	75 %
1345678	01132	01210	8	6	75 %
1783456	01132	01210	8	6	75 %
2348900	01132	01210	8	6	75 %
2784568	01132	01210	8	6	75 %
3456600	01132	01640	8	1	12.5 %
3722399	01132	11020	8	1	12.5 %

Notes: N^{WZ03} and N^{WZ08} refer to the (stylized) number of firms in a specific sector in 2003 and 2008, respectively. For a given sectoral combination, Max(share) is the share of the firms that moved to the converted sector (WZ08) as a fraction of the total number of firms in the previous sector (WZ03).

The original dataset, before any data treatment, contains 6.51 million observations for 2007 and 6.76 million observations for 2008. After the first phase of the data preparation process, which drops missing observations in relevant variables and focuses on firms, the dataset contains 10.3 million observations.⁶ About

⁶ The number of observations refers to the dataset used for creating the five-digit industry conversion table. The number of observations change with the level of detail in industry codes.

1.8 million of these firms do not have a match in either of the two years. As either the old or the new industry codes is not observed, these firms are also dropped from the dataset, leaving a final dataset of about 8.5 million observations, i. e. 4.25 million firms per year.

3 Conversion tables and robustness checks

We create conversion tables – using the methodology described in Section 2 – for the five-digit, four-digit, and three-digit codes. For test purposes, we also create a conversion table at the two-digit industry level, but it is not part of the published tables. The estimation process is applied for both directions, i. e. conversion from WZ03 to WZ08 and vice versa. The resulting conversion tables are provided as Excel files.⁷

The vast majority of industry pairs is defined by very high shares of firms assigned to the respective old (new) industries. Table 6 in the appendix displays summary statistics of the shares of firms that determine which single WZ03 sector a WZ08 code is assigned to (or vice versa). The table reveals that these shares are very high on average, i. e. roughly between 85 and 90 %. Even the 5th percentile of the shares is usually around 50 % or higher.⁸

In rare cases, however, the share of companies determining the conversion can be as low as 20 %. This reveals the main shortcoming of our approach, namely that sometimes less than half of the firms of a WZ03 code may determine the single WZ08 sector to which all firms from one WZ03 code are assigned to (see minimum values in the sixth column in Table 6). On the other hand, this is a common problem with conversion tables and only holds for a small fraction of our converted industries.⁹ The conversion tables include the information on the

Hence, the final dataset is larger if the three-digit industry level is considered relative to the four-digit or the five-digit industry level.

⁷ The tables are available for download from <https://bit.ly/2MPPhBL8> (DOI:dx.doi.org/10.25652/diw_data_S0018.1).

⁸ Moreover, Figure 1 in the appendix graphs the shares of companies determining the converted sector for the case of the four-digit conversion table, both for the direction WZ03 to WZ08 and vice versa. It can be seen that for most sectors, this share is very high (i. e. close to 100 %). The kernel density plots for the other conversion tables (five-digit level and three-digit level, respectively) look qualitatively similar.

⁹ For most of the six conversion tables, around five percent of the industry pairs are determined by less than 50 % of firms in a given industry. The average fraction determining the converted sector, however, is often close to 90 % (see Table 6).

respective share for each industry pair. Table 2 illustrates the benefits of using a more refined conversion table as industries are matched more accurately.

Example. When looking at the 4-digit level in Table 2, all manufacturers of photographic chemical material (WZ03 code D24.64) are assigned to industry C20.59 (WZ08 code), which is part of the industry manufacturing of chemicals and chemical products (WZ08 code C20). Firms producing prepared unrecorded media, on the other hand, end up in the sector manufacture of computer, electronic and optical products (WZ08 code C26). If only the three-digit level is available, firms from both industries would be seen as belonging to one single sector prior to 2008 (WZ03 code D24.6). Applying the conversion table for 3-digit industry codes, manufacturers of prepared unrecorded media (WZ03 code D24.65) would end up in the chemical industry (WZ08 code C20) instead of the correct industry (Manufacture of computer, electronic and optical products, WZ08 code C26).

Table 2: Differences arising from the use of the four-digit and the three-digit industry codes.

WZ03 (pre-2008), 4-digit	WZ03 (pre-2008), 3-digit	Name of industry (4-digit)	WZ08 (post-2008), according to 4-digit FSO	WZ08 (post-2008), according to 4-digit CT	WZ08 (post-2008), according to 3-digit CT
D24.64	D24.6	Manufacture of photographic chemical material	C20.59	C20.59	C20.5
D24.65	D24.6	Manufacture of prepared unrecorded media	C26.80	C26.80	C20.5

Source: Statistisches Bundesamt (2008b).
Notes: CT is the acronym for Conversion Table; FSO is the acronym for Federal Statistical Office.

In order to assess the relevance of the issue of misallocated industries, we check the differences between using the three-industry conversion table compared to a four-digit conversion tables, the differences between the four-industry conversion table and the five-digit conversion tables etc. We count the number of sectors that end up in the wrong industry once a less detailed conversion table is applied. Table 3 reveals that a significant share of the sectors is misallocated with a decreasing level of precision. For example, when using the three-digit

Table 3: Share of misallocated industries.

		In comparison with conversion table for			
		2-digits	3-digits	4-digits	5-digits
Base conversion table	2-digits	–	–	–	–
	3-digits	20.91 %	–	–	–
	4-digits	18.79 %	12.13 %	–	–
	5-digits	20.95 %	17.08 %	12.07 %	–

industry conversion table instead of the four-digit industry conversion table, 62 industries or 12.1 % of all four-digit sectors are attributed to the wrong three-digit industries. In total, more than 600,000 firms or about 13.7 % of the observations are affected. As can be seen from Table 4, in general, more industries are misallocated, the less detailed is the conversion table.

Next, we test the effects of applying conversion tables at different levels of detail on investment statistics from the *AFiD-Panel Manufacturing Firms*. These data are the base for capital stock estimations and, thus, relevant for micro- and macroeconomic productivity estimations or investment studies. The original data come with four-digit industry codes. However, deflator time series are often only available at the two-digit industry level. Moreover, most firm level studies use the two-digit level to either create industry dummies or to conduct estimations separately for every two-digit industry. Therefore, we apply the industry conversion tables for the two-digit and the four-digit industry codes and the years 2003 to 2007, checking the consequences for the investment data at the two-digit industry level.

Table 4 shows the effects of applying conversion tables for industry codes at the four-digit level versus the two-digit level on firm-level data that contain the old industry codes (WZ03). The table contains the investment sums according to the new industry classification system (WZ08), the number of observations in each industry, and the differences coming from using the two different conversion tables.

The first thing to notice is the fact that some manufacturing industries have no observations when the two-digit industry code is used and the respective conversion table applied. This is the case for industries B5, C11, C21, and C33.¹⁰

¹⁰ This is the result of dissolving old industries and creating new industries due to the change from WZ03 to WZ08. This has consequences at the two-digit industry level. For example, the pharmaceutical industry (D244) was part of the chemical industry (D24) in WZ03. It is now an own two-digit industry (C21). If the conversion is conducted at the two-digit level, all firms from D24 are assigned to C20 (Chemical industry) instead of C21.

Table 4: Differences in gross investment with different conversion tables, 2003–2007, WZ08.

Industry WZ08	Using 2-digit conversion table		Using 4-digit conversion table		Differences in investment		
	Total investment [†]	N	Total investment [†]	N	N	Total [†]	Mean ^{†,‡}
B5	–	–	2,001.64	48	48	2,001.64	41.70***
B8	3,735.13	1,917	1,733.49	1,869	–48	–2,001.64	–1.02***
B9	1,677.15	90	1,677.15	90	0	0.00	0.00
C10	21,675.40	24,734	15,078.81	22,016	–2,718	–6,596.59	–0.19***
C11	–	–	6,596.59	2,718	2,718	6,596.59	2.43***
C12	861.92	117	861.92	117	0	0.00	0.00
C13	2,143.74	4,195	2,069.56	3,873	–322	–74.18	0.02
C14	501.24	1,911	575.42	2,233	322	74.18	0.00
C15	345.35	872	345.35	872	0	0.00	0.00
C16	3,344.50	5,070	3,344.50	5,070	0	0.00	0.00
C17	8,527.26	4,119	8,527.26	4,119	0	0.00	0.00
C18	6,123.72	12,305	4,350.53	7,827	–4,478	–1,773.19	0.06**
C19	5,055.15	246	4,718.44	229	–17	–336.72	0.06
C20	29,195.68	6,917	21,009.54	5,568	–1,349	–8,186.14	–0.45
C21	–	–	8,487.42	1,339	1,339	8,487.42	6.34***
C22	11,616.75	13,146	11,616.75	13,146	0	0.00	0.00
C23	7,937.51	7,967	7,937.51	7,967	0	0.00	0.00
C24	11,967.68	4,505	11,967.68	4,505	0	0.00	0.00
C25	16,507.16	30,682	16,561.37	29,902	–780	54.21	0.02
C26	11,636.21	3,531	13,584.26	7,485	3,954	1,948.05	–1.48***
C27	12,245.58	9,606	12,346.29	9,498	–108	100.71	0.03
C28	24,431.79	29,789	21,796.43	28,545	–1,244	–2,635.37	–0.06
C29	54,639.75	4,926	56,521.49	5,482	556	1,881.74	–0.78
C30	5,032.07	1,548	4,189.39	742	–806	–842.68	2.40*
C31	3,311.34	7,539	1,940.16	5,088	–2,451	–1,371.17	–0.06*
C32	5,125.51	10,111	4,153.26	8,002	–2,109	–972.26	0.01
C33	–	–	1,872.21	3,015	3,015	1,872.21	0.62***
E38	668.64	806	668.64	806	0	0.00	0.00
J58	–	–	1,760.90	4,420	4,420	1,760.90	0.40***
J59	–	–	12.30	58	58	12.30	0.21***

Source: DOI:10.21242/42111.2014.00.01.1.1.0, own calculations.

Notes: *** significant at the 1 % level, ** significant at the 5 % level, * significant at the 10 % level, [†] in million Euros, [‡]significance based on two-sided t-test of differences in means (Welch's t-test).

This highlights the fact that the industry codes cannot be meaningfully converted at this level of detail, which is why we do not provide these conversion tables. Secondly, some firms end up in industries outside manufacturing when the four-

digit industry code and the respective conversion table is used. This is the result of the significant reorganization of industries in WZ08 compared to WZ03.

Example. Firms belonging to industry D22 (Publishing, printing and reproduction of recorded media) according to WZ03, are now largely part of industry J58 (Publishing activities) in WZ08. This means they are considered services in the new classification system.

The third fact to acknowledge is that the differences in investment are sometimes quite large, even when not significant according to the two-sided t-test of differences in means (last column Table 4). A good example is *manufacture of machinery and equipment* (C28), where the difference in nominal investment is 2.6 billion Euros.

To finally check the validity of our results, we compare our conversion tables with the conversion tables of the Federal Statistical Office of Germany (Statistisches Bundesamt 2008b). These latter tables cannot be directly used in applied research, since they simply contain all possible combinations between old and new sectors, without specifying what fraction of which sector in WZ03 was assigned to which sector in WZ08.¹¹ Since no further information is provided, there is no way of deciding which of the pairs should be used. However, we use these tables to test whether our conversion tables only contain combinations of old and new industry codes that are also listed in the tables of the Statistical Office.

Virtually all of the 1,037 combinations in our conversion table at the five-digit level also exist in the conversion table of the Federal Statistical Office of Germany. The only exceptions are listed in Table 5, namely the industries 01413, 15920, 22122, 55236, 60214, 64110, and 72301. This deviation means that – provided the conversion table of the Federal Statistical Office does not contain mistakes – either the majority of firms in these seven sectors had the wrong industry classification in the 2007 *AFiD-Panel Business Register*, or that the majority of these firms switched to another industry in the administrative process of assigning the new industry codes. A detailed analysis of the data reveals that, while the majority of firms in 2008 has an industry code that is listed in the fourth column of Table 5, indeed very few firms exist with the industry code that is listed in the Federal Statistical Office of Germany's conversion table (WZ08^{FSO} in Table 5). In some cases, less than three observations have the industry code combination that the conversion table of the Federal

¹¹ Using the example from Table 1, this conversion table would contain the pairs 01132–01210, 01132–01640 and 01132–11020.

Table 5: Deviations from the conversion table of the Federal Statistical Office of Germany.

WZ03	Name	N ^{WZ03}	WZ08 ^{CT}	Max (Share)	Sectors listed in WZ08 ^{FSO}
01413	Laying out, planting and maintenance of graves	3839	96032	86.1 %	81309
15920	Production of ethyl alcohol from fermented materials	464	11010	92.7 %	20140
22122	Publishing of weekly and Sunday newspapers	654	58130	79.1 %	58140
55236	Boarding houses	83	55909	84.3 %	55102 55101
60214	Operation of funicular railways and aerial cable-ways	290	93290	91.0 %	49310 49399
64110	National post activities	575	53200	71.7 %	82190 53100
72301	Data entry services	1122	62030	74.7 %	74201 63110

Source: DOI:10.21242/52111.2012.00.01.1.1.0, Statistisches Bundesamt (2008b), own calculations.

Notes: N^{WZ03} refers to the number of firms in a given five-digit WZ03 sector in the AFID-Panel Business Register in 2007. WZ08^{CT} is the five-digit WZ08 sector determined in our conversion tables. Max(Share) is the share of the firms that determines this converted WZ08 sector. WZ08^{FSO} are the sectors that are listed in the conversion tables of the Federal Statistical Office of Germany.

Statistical Office suggests. Consequently, the share of companies in these WZ08 sectors is typically very low.

As seen in Table 5, the total number of companies in industries that are missing in the Federal Statistical Office of Germany's conversion table is quite low. A total of 7,027 firms belong to these seven sectors, which is less than 0.02 % of all companies in the 2007 AFID-Panel Business Register. We conclude that, for the overwhelming majority of industry conversions, there is a concordance between our conversion table and the one from the Federal Statistical Office of Germany.

4 Summary

This study provides ready-to-use conversion tables for industry codes from the German industry classification system WZ 2003 to WZ 2008 (and vice versa) at the

five-digit level, the four-digit level, and the three-digit level of industry codes. These conversion tables are constructed using both the 2007 and 2008 German Business Registers, which contains observations for more than four million firms in each year. Although initially developed to be applied to firm-level data from the Statistical Offices, the tables can be applied to all German firm-level data that contain industry codes at the three-digit industry level or higher.

The final conversion tables contain unique pairs of industry codes to allow for their application. Each pair is defined by the relative majority of firms that moved from an old industry (WZ03) to a new industry (WZ08) or vice versa. This approach comes at a cost because some of the old industries have been completely split up. However, the statistics we provide in this study show that the overwhelming majority of industry pairs is defined by very high shares of firms assigned to the respective old (new) industries.

This paper further shows that the level of detail of the available industry codes affects the quality of the conversion. The tests reveal that the conversion works best at the highest level of detail of the original industry code. A final robustness check confirms that, with very few exceptions, the industry pairs in the conversion tables are consistent with the list of possible industry pairs that is published by the German Statistical Office. This latter list cannot be directly used in applied research, as it contains all potential industry pairs without specifying relative majorities. In contrast, direct applicability is the major advantage of the conversion tables we propose in this paper.

Acknowledgements: We thank Caroline Stiel for helpful suggestions. This research was partly funded by the German Ministry of Economics and Energy through the project: “Wissensbasiertes Kapital in Deutschland: Analysen zu Produktivitäts- und Wachstumseffekten und Erstellung eines Indikatorsystems.”

References

- Fritsch, M., B. Götzig, O. Hennchen, A. Stephan (2004), Cost Structure Surveys for Germany. *Schmollers Jahrbuch / Journal of Applied Social Science Studies* 124: 557–566.
- Greulich, M. (2009), Revidierte Wirtschaftszweig- und Güterklassifikationen fertiggestellt. *Wirtschaft und Statistik* 1: 36–46.
- Konold, M. (2007), New Possibilities for Economic Research through Integration of Establishment-level Panel Data of German Official Statistics. *Schmollers Jahrbuch / Journal of Applied Social Science Studies* 127: 321–334.
- Malchin, A., R. Voshage (2009), Official Firm Data for Germany. *Schmollers Jahrbuch / Journal of Applied Social Science Studies* 129: 501–513.

- Statistische Ämter des Bundes und der Länder (2015), Unternehmensregister – System 95. Metadaten für die On-Site-Nutzung. Stand 10.02.2015, zuletzt aufgerufen am 2019-08-07: https://www.forschungsdatenzentrum.de/sites/default/files/urs_95_mdr.pdf.
- Statistisches Bundesamt (2008a), Klassifikation der Wirtschaftszweige mit Erläuterungen – 2008. Stand 12.2008, zuletzt aufgerufen am 2019-08-07: <https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/klassifikation-wz-2008.html>.
- Statistisches Bundesamt (2008b), Umsteigeschlüssel der Klassifikation der Wirtschaftszweige, Ausgabe 2003 (WZ 2003) zur Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008). Stand 20.11.2008, zuletzt aufgerufen am 2019-08-07: <https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/Downloads/Klassifikationen-wz-2008-umsteiger.xls.html>.
- Statistisches Bundesamt (2016), Kostenstrukturerhebung im Verarbeitenden Gewerbe sowie des Bergbaus und der Gewinnung von Steinen und Erden. Stand 26. 05.2015, zuletzt aufgerufen am 2016-07-15: <https://www.destatis.de/DE/Publikationen/Qualitaetsberichte/Qualitaetsberichte.html>.
- Vogel, A. (2009), The German Business Services Statistics Panel 2003 to 2007. *Schmollers Jahrbuch/Journal of Applied Social Science Studies* 129: 515–522.
- Wagner, J. (2000), Firm Panel Data from German Official Statistics. *Schmollers Jahrbuch/Journal of Applied Social Science Studies* 120: 143–150.
- Wagner, J. (2008), Die Forschungspotentiale der Betriebspaneldaten des Monatsberichtes im Verarbeitenden Gewerbe. *ASTa – Wirtschafts- und sozialstatistisches Archiv* 2: 209–221.
- Wagner, J. (2010), The Research Potential of New Types of Enterprise Data Based on Surveys from Official Statistics in Germany. *Schmollers Jahrbuch/Journal of Applied Social Science Studies* 130: 133–142.
- Zühlke, S., M. Zwick, S. Scharnhorst, T. Wende (2004), The Research Data Centres of the Federal Statistical Office and the Statistical Offices of the Länder. *Journal of Applied Social Science Studies* 124: 567–578.

Appendix

Table 6: Shares of the companies determining the converted industry, by direction of conversion.

Level of precision	Direction of conversion	Mean	5th percentile	95th percentile	Minimum
Five-digit	WZ03 → WZ08	92.2 %	70.1 %	1	33.3 %
Five-digit	WZ08 → WZ03	82.6 %	41.5 %	98.8 %	20.0 %
Four-digit	WZ03 → WZ08	86.6 %	56.1 %	99.6 %	31.3 %
Four-digit	WZ08 → WZ03	86.2 %	46.8 %	98.9 %	19.3 %
Three-digit	WZ03 → WZ08	82.9 %	48.1 %	99.2 %	23.1 %
Three-digit	WZ08 → WZ03	85.9 %	49.9 %	99.0 %	19.3 %

Source: DOI:10.21242/52111.2012.00.01.1.1.0, own calculations.

Figure 1: Kernel densities of the shares of companies determining the converted industry (four-digit).

Source: DOI:10.21242/52111.2012.00.01.1.1.0, own calculations.