RESEARCH DATA CENTRE (FDZ)
of the German Federal Employment Agency (BA)
at the Institute for Employment Research (IAB)

FDZ-METHODENREPORT

Methodological aspects of labour market data

06|2025 EN Linking the Mannheim Enterprise Panel (MUP) with Administrative Establishment Data of IAB – an update

André Diegmann, Thorsten Doherr, Mirja Hälbig, Stefanie Wolter

Bundesagentur für Arbeit

Linking the Mannheim Enterprise Panel (MUP) with Administrative Establishment Data of IAB – an update

André Diegmann (Institute for Employment Research, Halle Institute for Economic Research, CESifo), Thorsten Doherr (Leibniz Centre for European Economic Research), Mirja Hälbig (Halle Institute for Economic Research), Stefanie Wolter (Institute for Employment Research)

Die FDZ-Methodenreporte befassen sich mit den methodischen Aspekten der Daten des FDZ und helfen somit Nutzerinnen und Nutzern bei der Analyse der Daten. Nutzerinnen und Nutzer können hierzu in dieser Reihe zitationsfähig publizieren und stellen sich der öffentlichen Diskussion.

FDZ-Methodenreporte (FDZ method reports) deal with methodical aspects of FDZ data and help users in the analysis of these data. In addition, users can publish their results in a citable manner and present them for public discussion.

Content

List of Figures

List of Tables

Zusammenfassung

Dieser Bericht beschreibt die Aktualisierung des Schlüssel zwischen Betriebs- und Unternehmensidentifikatoren. Er ermöglicht die Verknüpfung von administrativen Beschäftigungsdaten des Instituts für Arbeitsmarkt- und Berufsforschung (IAB) mit externen Unternehmensdaten für die Jahre 1993 bis 2022, die über einen Creditreform- oder Moody's (ehemals Bureau van Dijk)-Identifikator verfügen. Die Zuordnung zwischen den beiden Datenquellen basiert auf einem Record-Linkage zwischen den Adressdaten des IAB und den Adressdaten des Mannheimer Unternehmenspanels (MUP) des Leibniz-Zentrum für Europäische Wirtschaftsforschung (ZEW). Dieser Methodenbericht fasst den Prozess der Datenverknüpfung zusammen, beschreibt ausführlich die Bereinigung der Daten und Qualitäts- und Repräsentativitätsprüfungen sowie Einschränkungen der Daten. Der Schlüssel wird die Grundlage für Standarddatenprodukte bilden, die über das Forschungsdatenzentrum des IAB (FDZ) verfügbar sein werden. Darüber hinaus ermöglicht der vorgestellte Schlüssel die Erstellung individueller Datensätze.

Abstract

This report describes the update of the key between establishment and enterprise identifiers. This key enables the combination of administrative employment data of the Institute for Employment Research (IAB) with external enterprise data that has a Creditreform or Moodys (former Bureau van Dijk) identifier for the years 1993 to 2022. To establish this combination, we performed a record linkage between the address data of the IAB and the address data of the Mannheim Enterprise Panel (MUP) hosted at the Leibniz Centre for European Economic Research (ZEW). This report describes the record linkage process, present quality and representativeness checks and also discusses limitations of the data. The key will be the base of standard data products that will be available via the Research Data Centre of the IAB (FDZ). Moreover, the presented key allows to generate customized data sets.

Keywords

Establishments, Enterprises, Establishment-History-Panel (BHP), Mannheim Enterprise Panel (MUP), MUP-BHP, Record-Linkage

Acknowledgements

We are grateful to Alexandra Schmucker for helpful comments that have improved this report and to the team of the ZEW, especially Sandra Gottschalk, for providing the Mannheim Enterprise Panel and their expertise for this linkage. Without the ZEW this would not have been possible.

1 Introduction

The analysis of worker and firm dynamics in Germany has been primarily focused on employees subject to social security employment within and across establishments. A large and pressing number of questions in labour economics and beyond require the knowledge of the enterprise identifiers behind establishments. Such enterprise IDs will allow researchers to better understand driving factors and interrelations between agents in the labour market. The IAB is known for high quality labour market data on employees and establishments but it is lacking information on how establishments cluster into enterprises or interrelate with each other. Moreover, the IAB has, aside from survey information, no data on firms' finances and balance sheet information. Several attempts were made to overcome this shortcoming by linking the IAB establishment level data with enterprise level information. However, previous data initiatives had a couple of shortcomings including (i) selectivity of the sample, (ii) limited data access and/or (ii) high costs (Antoni et al. 2018, Bender et al. 2007). These shortcomings are addressed with the new link between the Mannheim Enterprise Panel (Mannheimer Unternehmenspanel, MUP) developed by the Leibniz Centre of European Economic Research (ZEW) and the administrative establishment information from IAB. We were able to create this data set in a collaboration with the ZEW. The central data of the MUP are originally collected by Creditreform e.V., a credit rating agency that is sharing the data with the ZEW for data quality reasons and economic research (https://www.zew.de/forschung/mannheimer-unternehmenspanel). This link overcomes the difficulties, enabling us to provide and offer for the first time an establishments-enterprise-level data set that covers the majority of the IAB establishments, offers a long time series of observations and will be part of Research Data Center of the Federal Employment Agency in the Institute for Employment Research's (hence force FDZ) standard data portfolio.

The new dataset not only makes it possible to answer new research questions, it also provides an interface to link external company information to administrative IAB data more easily in the future. The reason is that the enterprise identifier in the MUP data coincides with common identifiers provided by Moody's Orbis data (https://www.moodys.com/web/en/us/capabilities/company-reference-data/orbis.html).

The aim of this report is to provide a documentation on the establishment enterprise key and to show the strength and limitations of the product. The remainder is structured as follows: Section 2 briefly defines the observation units. Section 3 summarizes the linkages process, followed by the steps to enhance data quality discussed in Section 4. Section 5 provides analyses on the representativeness of the key. Finally, Section 6 contains information on data access.

2 Identifying establishments and enterprises

Working with establishment and firm level data is challenging as different levels of economic units exist for different legal purposes. Not all units can be mapped perfectly into each other. The data

we focus on maps establishments to legal units or enterprises. For simplicity, we follow the naming convention introduced by the MUP and call the legal units enterprises. This chapter shortly summarizes the differences between the units used throughout the report.

2.1 Establishments

The data report on the Establishment History Panel (BHP) defines what an establishment is (Ganzer et al. 2023). We summarize the definition here and especially point out particularities that are relevant for working with linked establishment and firm data. An establishment is a regionally and economically delimited unit in which employees work. An establishment may consist of one or more branch offices or workplaces belonging to one enterprise. The eight-digit establishment identifier is issued by the Federal Employment Agency. It identifies a unit uniquely and is issued if an establishment is obliged to submit social security notifications. Not in all cases an establishment identifies a physical unit, e.g., a production plant or an office as the assignment depends on the economic purpose of the establishment and its main economic activity. Larger production sites therefore could hold multiple identifiers if the site hosts more than one economic purpose, e.g. research and development, production and catering. An establishment can be economically independent but it does not need to be. Hence, it reports employees to the Federal Employment Agency but, for instance, no information to the tax authorities.

The establishment identifier is time consistent and unaltered usually when economic activities or the address change, yet there are events that result in a change of the identifier. Reasons for the change in the identifier can be:

- Change in the ownership. However, if all parties agree the establishment identifier can remain unchanged
- A merger and acquisition activity or if the company is split up
- A move to another municipality where an establishment of the same company exists. Then only one establishment number survives.

More details on how establishment identifiers are allocated can be found in Ganzer et al. (2023) and in the documentation by the BA Employer service (2022).

2.2 Legal Unit

“A legal unit in the statistical business register is a natural person who is economically active, a legal person or an association of persons. Economic activities also include the exercise of a liberal profession and the holding of participations in other legal units. Thus, for example, a joint-stock company, limited liability company, general partnership or sole proprietor are considered to be legal units.”[1] A legal unit is the smallest unit that is obliged to prepare financial statements or similar records for commercial and/ or tax law. From this definition it becomes clear that a legal unit can consist of multiple establishments but not all legal units are establishments vis a vis production units.

[1] translated from https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Unternehmen/Glossar/rechtliche-einheiten.html

2.3 Enterprise

The EU introduced an additional definition to harmonize information across countries for statistical purposes and to reflect the nature of complex enterprises. The European definition considers that some independent legal units carry out activities only for one other legal unit, i.e. they act as sub-unit to the main enterprise. These enterprises were often split-off for tax-reasons or to avoid liabilities[2]. For example, a legal unit that provides catering to only one business customer and does not act as independent actor on the market would be considered as one enterprise jointly with the legal unit it provides catering for.

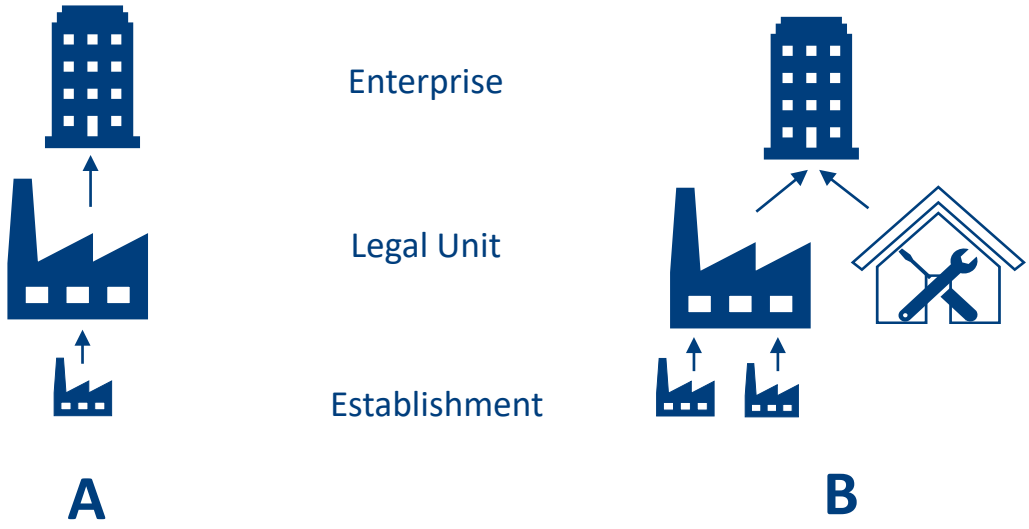
2.4 Enterprise group

Independent enterprises can cooperate in enterprise groups. These groups are often multinational and contribute a substantial share to the national economy (Ahlborn et al. 2021). We only have data for Germany. Thus, the mapping of establishments into multinational enterprises is not feasible with these data.

Figure 1 illustrates the relations between our definitions. The simple case is case A on the left-hand side. The factory symbol at the bottom represents an establishment. This might be a production plant that employs a certain number of employees liable to social security as well as some marginal part-time workers and is owned by a natural person. In this case, the establishment that reports the number of employees is the same legal unit that reports to the tax authorities as well as the unit that is defined as an enterprise in the sense of the EU definition. In this case, only one physical unit is observable. Case B visualizes a complex enterprise with two legal units. While the left legal unit in case B contains two establishments, the legal unit on the right side is not associated with an establishment in the IAB data. Moreover, the legal units are both economically active followong the definition of the German Federal Statistical Office but both are part of the same enterprise. Establishments in case B could be production plants in two different municipalities, i.e., they have two different establishment identifiers. The legal unit without any establishment could represent a legal entity that owns machinery and lends them only to the establishments of the other legal unit in B. Since this legal unit is not independently active on the market, it would be recognized as one enterprise with the first legal unit.

[2] https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Enterprise

Figure 1 Illustration of establishments, legal units and enterprises



Source: own illustration.

3 Linking establishment and enterprise information

3.1 Addresses of the Mannheim Enterprise Panel

The MUP is a data base that is hosted at the Centre for European Economic Research (ZEW). The data contains biannual firm-level information collected by Creditreform e.V., the largest German credit rating agency, sent to ZEW for further data processing (Bersch et al. 2014).

The entity used in the MUP data closely corresponds to the legal units described in Section 2.2. with the restriction that Creditreform focuses on legal units with sufficient economic activities. This partly excludes freelancers, microenterprises, especially in the agricultural sector, and side-line businesses. However, it cannot be ruled out that there is some enterprise structure, as defined by the EU standards, embedded in MUP.

MUP is highly representative for enterprises in Germany. By comparing the MUP data to the Business register data of the Federal Statistical Office, Bersch et al. (2014) find that the MUP captures 90 percent of the Business register firms. The MUP sample is slightly biased towards larger firms and underrepresent sole proprietors (Bersch et al. 2014).

In order to link IAB establishments to MUP firms, we make use of the name and addresses information as well as information on economic activity.

3.2 Establishment addresses of IAB

Establishments with at least one marginal part-time worker have to apply for an identifier at the Federal Employment Agency's Employer service. They have to notify the establishment's name and address and the main economic activity. For the linkage we use the information on names and addresses to find matches in the firm level data. Note that not all establishments that register do later report employment. These establishment are kept in the linkage procedure.

As indicated at the beginning, the IAB data does not cover self-employed or establishments without employees liable to social security, e.g., civil servants. Hence, the intersection of both address sources is not identical.

The following units are represented in both data sources

- Employers in the private sector with at least one employee

The following units are partly or not at all included in the MUP data but in the IAB data

- Microenterprises
- Public administration
- Associations with employees
- Private households with household staff
- Agriculture, fishing and forest industry

The following units are not included in the IAB data but at least partly in the MUP data

- Freelancers
- Self-employed
- Firms with all employees not liable to social security

3.3 Record-Linkage Procedure

This chapter briefly summarizes the record linkage. A detailed and more technical description on the record linkage procedure can be found in Doherr (2023). In contrast to the first linkage (Diegmann et al., 2024), we allow for variation in establishments' names and addresses by year. We hope that this enables us to take better account of changes in crefo identifier for firms over time as well as to track changes of establishments assignment to different firm identifiers. The latter might hint to potential mergers and acquisitions, which we hope this updated link helps to shed light on in future research. Each variation is matched separately. The linkage includes address data for the years 1991 to 2022 for IAB data. The IAB address data for 1990ies is incomplete and of lower quality than subsequent years. The MUP data covers the years 1993 to 2022 and also has a patchy firm population representation of the 1990s. The reason is that the ZEW required merely start-ups, questionnaire samples and east German companies for research projects before 2000. Only since then, the contracts with the source data provider (creditreform) were changed to entail the full data coverage for the MUP.

We first undertake a pre-processing procedure, including the harmonization of enterprise names and dropping duplicates with respect to name, address, and economic activity. Adding address data for the longer time period has little impact on the overall number of establishments and enterprises that is available for the record linkage. On establishment side we extended the period from 2000-2018 to 1991-2022 and on the enterprise side from 1993-2020 to 1993-2022. We end up with 14.8 million establishments (previous link 14.3 million) and 13.8 million enterprises (previous link 13.3 million). The slight increase in establishments and firms relative to the large increase in years underlines the undercoverage for the period before 2000.

The record linkage was performed on the name of the entity, street, postcode and city. For the linkage process, the ZEW's Search Engine was used. In general, we put the focus on the similarity of establishment and enterprise names delegating a supportive role to the address fields. The linkage procedure runs in seven steps. In the first round, we impose the highest threshold with 85 percent enforcing a significant overlap of addresses. With every consecutive step, we relax the matching criteria for the respectively remaining unmatched cases until we use 3-grams for the names and renounce the necessity of any address similarity with a threshold of 65 percent in the last step. For detailed technical information see Doherr (2023). The Search Engine produces a link of 18 million establishments to over 60 million firm candidates. To classify matches, a Neural Network is trained with a small training sample of 2,000 establishments and 5,523 enterprise candidates that were manually classified in matches and non-matches. As the raw output before further cleaning steps, the linkage procedure provides a data set that links 8,277,594 (old: 8,273,283) establishment identifiers to7,339,994 (old: 7,281,447) firm identifiers. Some establishment – enterprise matches from the old key are no longer classified as matches and vis versa.

4 Enhancing the quality of the key

The key data combines the establishment and enterprise identifier into one data set. The key, linking establishments to firms, contains duplicates in terms of establishments that are assigned to multiple firms (at the same point in time). To enhance the quality of the record linkage key, we focus on establishments that we can link to the BHP between 1990 to 2022. The BHP is an ideal starting point for further data cleaning as it includes the overwhelming majority of establishments that are active in the labour market. Only establishments that did not have at least one marginal employee on the cut-off date of 30 June or December 31st are not represented. Our goal is to construct a panel data that assigns an establishment a unique enterprise identifier for each year from 1990 onwards. After linking the key to the BHP data, we are left with 6,536,340 linkable establishments of which 80 percent match to exactly one crefo ID. Overall we find at least one match in the BHP for 6,348,856 crefo IDs. The main reason for the reduced number of establishments compared to the raw link is because of zero employment numbers. The main remaining challenge is to select the most suitable enterprise match when two or more enterprises are potentially good matches, considering name and location changes. We have 5,223,340 establishments that link to exactly one crefo and another 351,627 establishments that linked to exactly one crefo for each name-address variation. For the remaining establishments, we have at least one name-address combination that is linked to multiple crefo-IDs. To improve the key, we conduct six cleaning steps.

Overlap of active years – While we match crefo IDs based on names and addresses of establishments at specific points in time, the resulting key has no explicit time structure; meaning that we observe links from establishment to crefo IDs that were not active at the same time. We add information from MUP on when the enterprise was active and from the BHP on first and last appearance of the establishment with reporting of social security employees. In this step, we remove 241,168 links where the active periods do not overlap. After this cleaning step 89,283 establishments or name-address variations of establishments are linked to exactly one crefo ID.

Match in locations – In the next step, we give preference to the link where the location between the establishment and the enterprise align. Table 1 provides an example. Establishment 1 has relocated from ZIP code 04229 in 2000 to ZIP code 01667 in 2007. There are two potential crefo IDs for establishment 1, one which is located in ZIP code 04229 (crefo ID 22) and one which is located in ZIP code 01667 (crefo ID 37). While leaving the establishment ID unchanged, we conclude that the change in location led to a change in the crefo identifier. We therefore select crefo ID 22 from 2000 onwards and crefo ID 37 from 2007 onwards. Establishment 2 has only one address line and is matched to two crefo IDs, one of which is in the same location. In that case, we select the crefo ID that matches in terms of location. This cleaning step provides us with another roughly 10,000 exact matches between establishment (and name-address variations) and crefo IDs.

Table 1 Fictional Example Location

Establishment ID	Postal code establishment	Year of establishment observation	Crefo ID	Postal code enterprise
1	**04229**	**2000**	**22**	**04229**
1	04229	2000	37	01167
1	01167	2007	22	04229
1	**01167**	**2007**	**37**	**01167**
2	04229	2003	45	04229
2	04229	2003	81	01167

Source: Own illustration

Reporting Information on employment – In this step, we drop links to crefo IDs which never report employment in the MUP if there is another crefo ID link with employment information. In this way, we favor links with information on employment in the MUP. This cleaning step creates another 265,859 exact combinations.

Overlap in employment and revenues reporting - We further extend step 4 by selecting crefo IDs for establishment name-address entries based on crefo reporting behavior from the MUP. We drop links where the crefo ID does not report employment or revenue values for the respective establishments' name-address validity period, if there is another link with overlap. Among links with overlap, we choose the one with the largest number of overlapping years. Table 2 provides an example. For the first entry of establishment 1 that is valid from 2000 until the updated entry in 2007, we choose the link with crefo ID 45 as it covers the years 2002 to 2006, while crefo ID 68 only covers the years 2002-2003. For the second entry of establishment 1 that is valid from 2007 onwards, we

choose the link with crefo ID 85, and drop the link with crefo ID 93 as the latter does not overlap with the date of the establishment entry. Table 3 shows the final link for this example. After this step we have 357,544 unique establishment-crefo combinations. For these remaining 218,843 we use information provided by the linking procedure to select the best possible match.

Table 2 Fictional Example for Selection based on Reporting Behavior.

Establishment ID	Year of establishment observation	Crefo ID	Employment values available in MUP from – to
1	2000	45	2002 – 2006
1	2000	68	2002 – 2003
1	2007	85	2007 – 2019
1	2007	93	2005 – 2006

Source: Own illustration

Table 3 Fictional Establishment - Crefo Panel

Establishment ID	year	Crefo ID
1	2000	68
1	2001	68
1	…	…
1	2006	68
1	2007	85
1	2008	85
1	…	…
1	2019	85

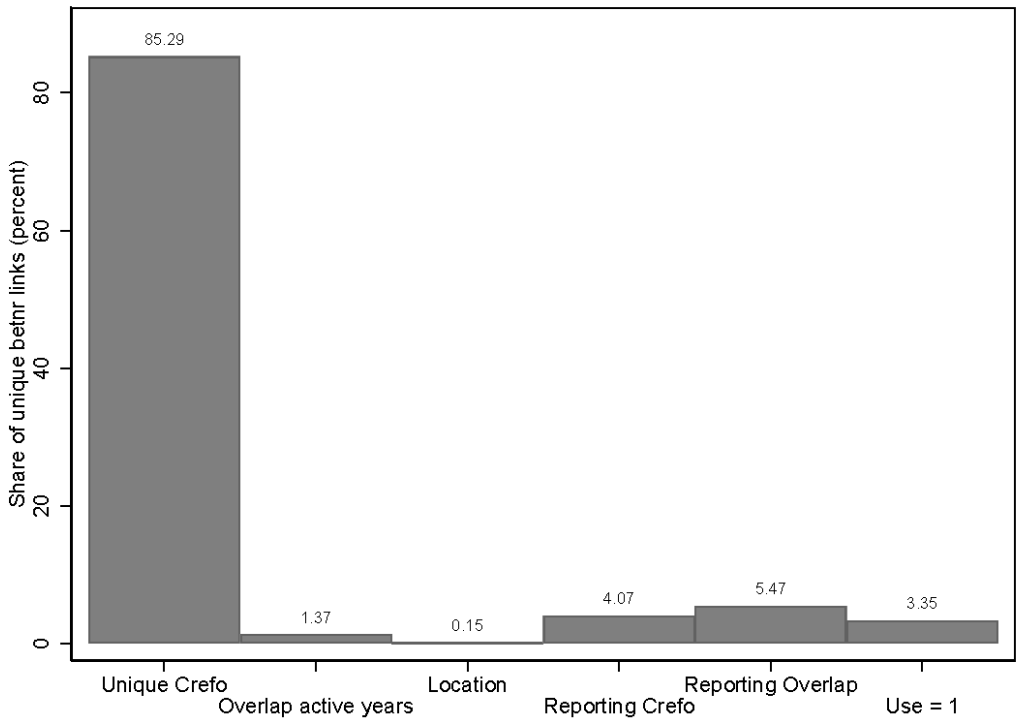
Source: Own illustration

Use recommendation -The Search Engine ranks the matches based on data quality, industry and geographical distance and provides a use recommendation for each establishment name-address entry. For the remaining multiple assignments, we follow this recommendation.

Figure 2 summarizes the cleaning steps and the share of unique betnr-crefo identifications we gain with each step. The final key contains an annual unique association between betnr and crefo ID for 6,516,438 unique establishments and with overall 77,902,307 establishment-year observations for the years 1990 to 2022[3]. We observe 2,334,380 establishments in 2000 and 2,566,286 establishments in 2020, and 2,105,686 enterprises in 2000 and 2,263,292 enterprises in 2020. In the final link, 90.99% of establishment are linked to one unique crefo, and 8.29% are linked to two different crefos between 1990 and 2022.

[3] We provide the link for the years 1990 to 2022. However, as the years before 1993 do not overlap with the record linkage the key is incomplete and biased for these years. Also, in the 1990ies and partly in the 2000ers the key is incomplete and biased. A good quality and completeness of enterprise information can be assumed from 2010 onwards.

Figure 2 Cleaning steps of the enterprise-establishment key



Source: MUP, BHP_1022_m06_v1, and key between MUP ID and betnr, own calculation.

5 Quality checks and representativeness

The linkage shows very high rates of overall quality and match rates. These match rates vary over time, industry and region. In this section, we first describe the match rates over these dimensions and provide a discussion on representativeness.

5.1 Link with Establishment History Panel

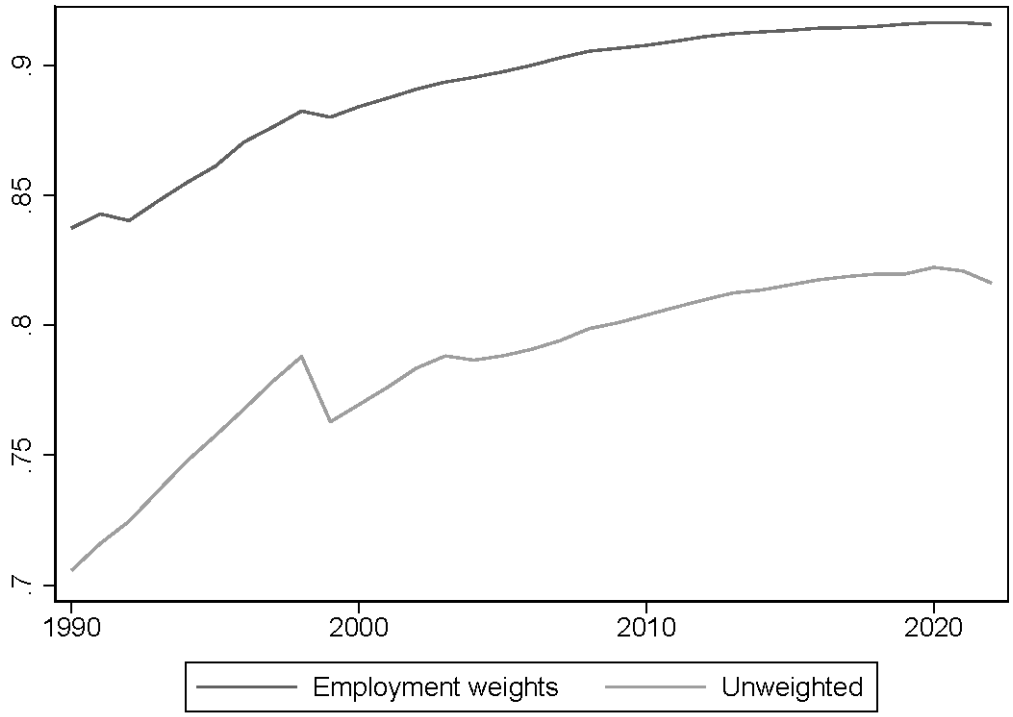
We are interested in the match rate based on all active establishments that have at least one employee reported during the years between 1990 and 2022. Thus, the following analysis does not consider establishments that do not show up in the BHP data files. BHP is composed of cross-sectional datasets since 1975 for West Germany and 1992 for East Germany. Every cross-section contains all the establishments in Germany which are covered by the IAB Employment History (BeH) on June 30th. These are all establishments with at least one employee liable to social security on the reference date. Establishments with no employee liable to social security but with at least one marginal part-time employee are included since 1999 (Ganzer et al. 2023). Based on this concept, the data contains up to 2.6 million linkable establishments and 2.1 million enterprises per year (Figure 3). The level and the trend are strongly reported to specific industries. As mentioned in Section 3.1, while IAB data has an employment concept in defining establishments, the MUP only lists economically active enterprises. To account for this, we exclude establishments from the BHP that are not economically active. For this reason, we further exclude establishments with a sector affiliation in public administration (industry classification 841), private households as employers (industry classification 970), goods production and service provision by private persons (industry classifications 981/982) and exterritorial organizations (industry classification 990). The vast majority of the excluded establishments are private households that act as employers. The dashed line shows the number of economically active establishments in the matched data.

Figure 3 Number of matched establishments and enterprises

Source: Years 1990 to 2022 of BHP_7522_m06_v1 and key between MUP ID and betnr, own calculation.

Figure 4 reports the share of matched economically active establishments and the share of employees that is covered by these matched establishments. The overall matching rate is around 81 percent of all establishment-year observations matched to an enterprise ID. The coverage of establishments increases over time from 79 percent to around 82 percent. These linked establishments account for over 90 percent of employment, which hints at differences in matching rates by firm size.

Figure 4 Share of covered establishments and employees

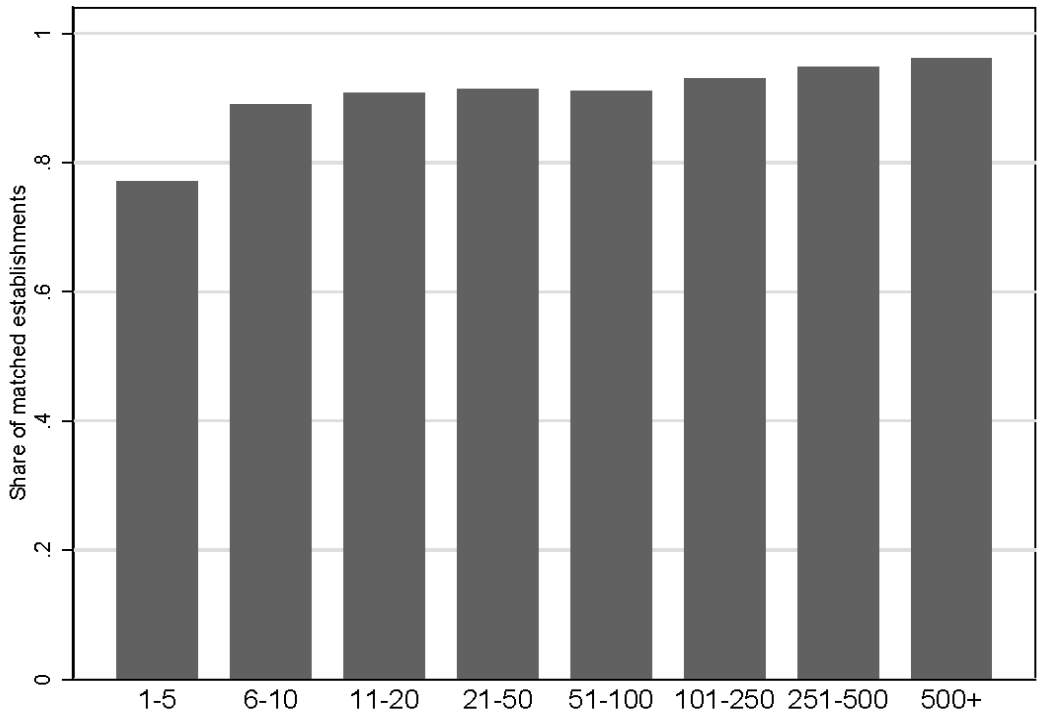


Source: Years 1990 to 2022 of BHP_7522_m06_v1 and key between MUP ID and betnr, excluding establishments that are economically inactive, own calculation.

We restrict our sample to the cross-section in 2019 to provide further information of the match rates with respect to firm size and industry. In the next step, we compare the match rates with regard to establishment size.

Figure 5 shows that the larger the establishment, the higher probability of being linked to an enterprise identifier. While we link three out of four establishments with less than six employees, the match rate is with 96 percent almost complete in the category of establishments with 500 and more employees.

Figure 5 Establishments that were assigned an enterprise identifier from the BHP 2019 by size

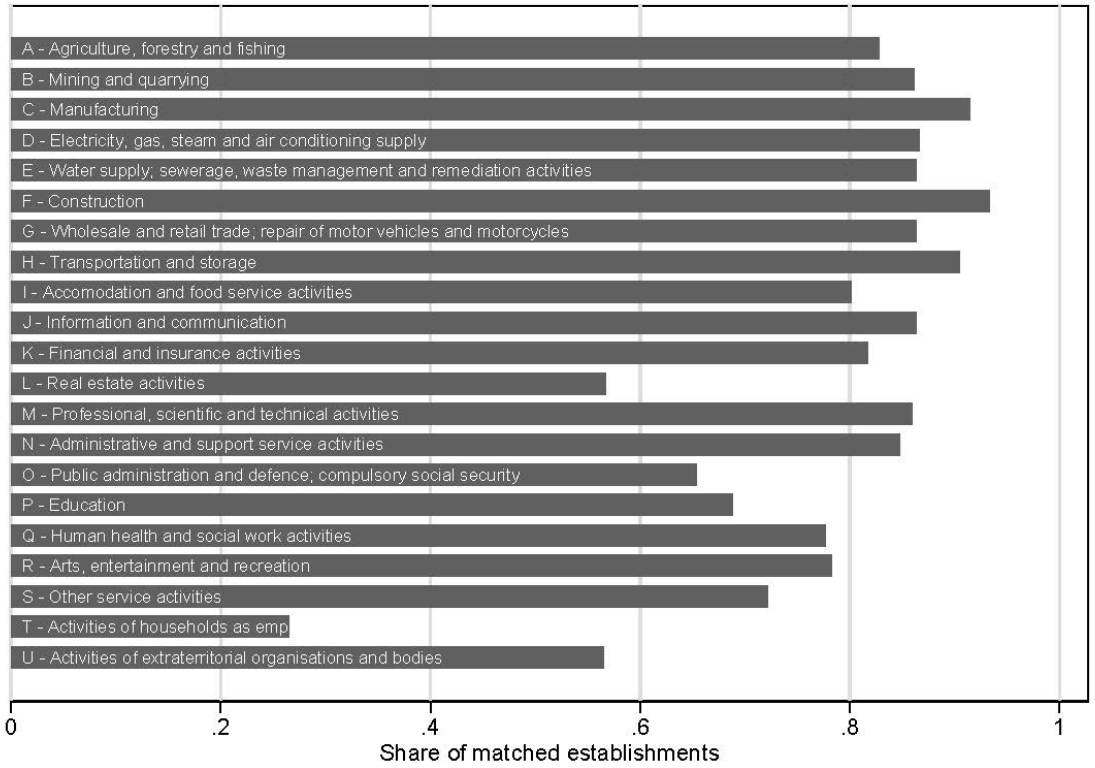


Source: Year 2019 of BHP_7522_m06_v1 and key between MUP ID and betnr, own calculation.

The distribution of establishments per enterprise is highly skewed. In our linked sample, 93.0 percent of all enterprises are single-establishment firms in 2000 which declines to 92.4 percent in 2020. Another six to seven percent of enterprises have two establishments. About 1.5 percent of all enterprises have three and more establishments. The number of single-establishment enterprises is comparable to other external sources for validation. The predecessor project that linked Orbis data with IAB data also found around 92 percent single-establishment enterprises (Antoni et al. 2018).

The linkage success also varies by industry. Industries with a high proportion of public establishments and sideline businesses have lower linkage rates compared to establishments in the construction and production sector. Figure 6 reports the match rates by economic sector.

Figure 6 Percent of Matched Establishments by Economics Sector (NACE Rev. 2), BHP 2019



Source: Year 2019 of BHP_7522_m06_v1 and key between MUP ID and betnr, own calculation.

5.2 Link with MUP

From the perspective of the MUP, the overall match rate amounts to 70 percent of the enterprises linked to an establishment in the BHP. This number also varies by industry and whether Creditreform contains employment information for the enterprise. The MUP provides information on the legal form, which is a further interesting margin to explore. Among firms for which employment information is available in the MUP, the match rate is with around 82 percent significantly higher. Table 4 shows the percent of firms from the MUP matched to an establishment from the BHP for selected years. The match rates decrease at the end of the sample period, because for the linkage IAB address data was only available until 2018.

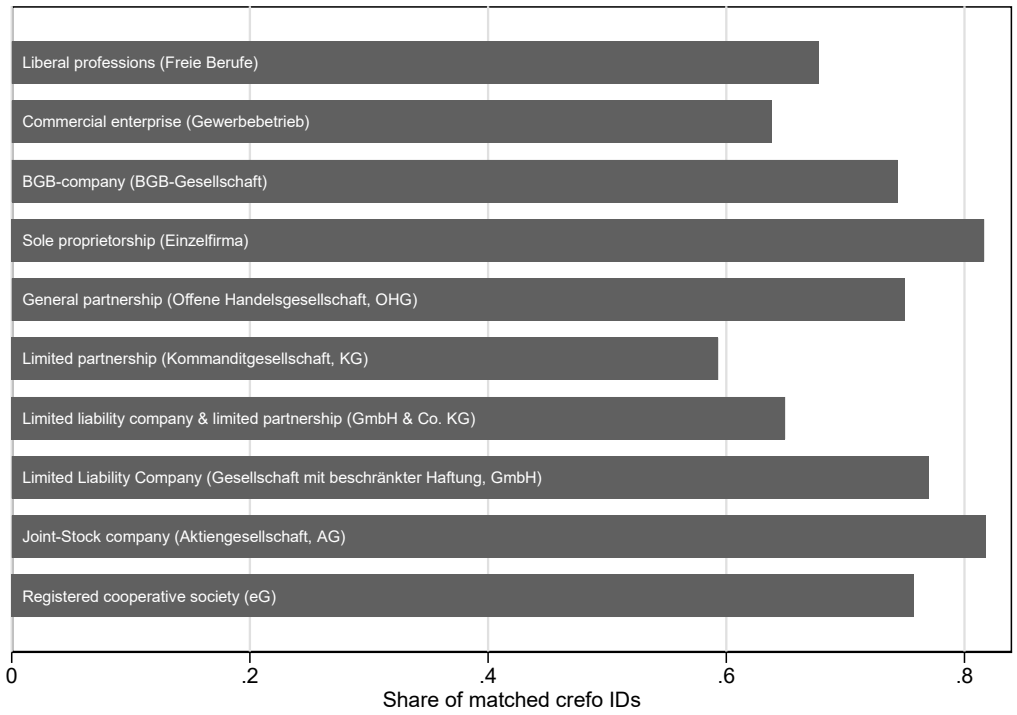
Table 4 Percent of MUP-Enterprises Matched to BHP Establishments.

Year	Percent Matched	Percent Matched (excl. firms without employment information)
2000	70.0	81.0
2005	70.0	81.0
2010	70.0	82.0
2015	71.0	84.0
2020	68.0	80.0

Source: Years 2000 to 2020 of MUP and key between MUP ID and betnr, own calculation.

There is considerable heterogeneity in match rates between legal forms. Figure 7 shows the match rates for the different legal forms covered in the MUP for the year 2019. While we can match around 77 percent of all limited liabilities (GmbH) and 82 percent of all stock corporations (AG), we find an establishment for 59 percent of all limited partnerships (KG) and 64 percent of all commercial enterprises.

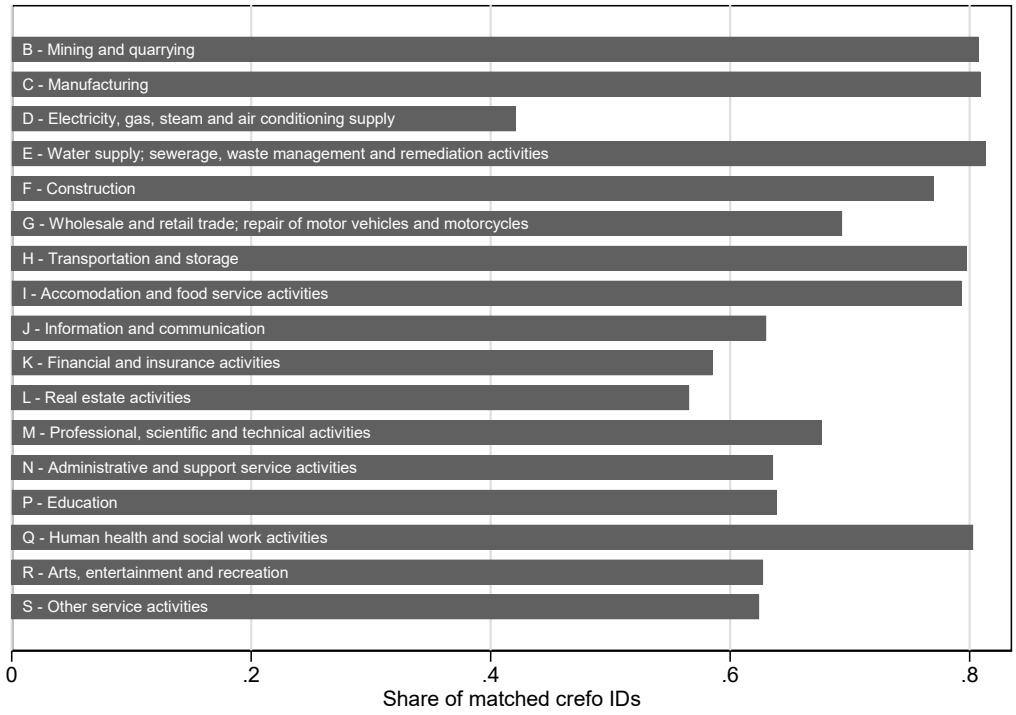
Figure 7 Percent of Matched Enterprises by Legal Form, MUP 2019



Source: Years 2019 of MUP and key between MUP ID and betnr, own calculation.

In addition, the match rate varies by economic sector. While we can match around 81 percent of all firms in the manufacturing sector, we find an establishment in only 57 percent of cases in financial and insurance activities. Figure 8 shows an overview over all economic sectors and the match rate with an establishment.

Figure 8 Percent of Matched Enterprises by Economic Sector (NACE Rev. 2), BHP 2019



Source: Year 2019 of MUP and key between MUP ID and betnr, own calculation .

6 Data access

The key between establishments and firms serves as base for further data products. Some of which the FDZ is and will develop as standard data sets. The first data product uses the link between establishments and firms to enrich the BHP with information from the Mannheim Enterprise Panel (MUP). This data is available as free standard data in the FDZ's data portfolio. For the subset of limited liability companies, the IAB-ZEW-MUP-BHP (DOI: 10.5164/IAB.MUP-BHP1023.de.en.v1) adds information on the shareholder structure and key financial information to the BHP data (Gottschalk et al. 2025). This data can be accessed for non-commercial labour market research after concluding a contract with the FDZ. Data access is possible via access in safe rooms and via the remote execution platform JoSuA. A second standard data set that combines MUP information and employee data from the IAB into a linked enterprise establishment employee dataset is under construction.

In addition to the standard FDZ data products, the key can be used to generate customized data linking external information that contains a Moody's or Crefo ID with administrative data from the IAB. For customized data the FDZ provides counseling on feasibility and costs. However, capacities for non-standard data products is limited. For more information on individualized data sets contact the German Record Linkage Center at recordlinkage@iab.de.

Literature

Antoni, Manfred; Koller, Katharina; Laible, Marie-Christine; Zimmermann, Florian (2018): Orbis-AD-IAB: From record linkage key to research dataset. Combining commercial company data with administrative employer-employee data. FDZ-Methodenreport 04/2028 (en), Nürnberg.

Bender, Stefan; Wagner, Joachim; Zwick, Markus (2007): KombiFiD – Kombinierte Firmendaten für Deutschland. Konzeption der Machbarkeitsstudie für eine Zusammenführung von Unternehmensdaten der Statistischen Ämter, des Instituts für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit und weiterer Datenproduzenten. IAB-Methodenreport 05/2007, Nürnberg.

Bersch, Johannes; Gottschalk, Sandra; Müller, Bettina; Niefert; Michaela (2014): The Mannheim Enterprise Panel (MUP) and Firm Statistics for Germany. ZEW Discussion Paper No. 14-104, Mannheim.

Bundesagentur für Arbeit (2022): Betriebsnummernvergabe. https://www.arbeitsagentur.de/datei/dok_ba031870.pdf, version: 12/21/2022, accessed: 06/26/2024, German only.

Diegmann, André; Doherr, Thorsten; Hälbig, Mirja; Schmucker, Alexandra; Wolter, Stefanie (2024): Linking the Mannheim Enterprise Panel (MUP) with Administrative Establishment Data of IAB. FDZ-Methodenreport, 03/2024 (en), Nürnberg, DOI: 10.5164/IAB.FDZM.2403.en.v1.

Gottschalk, Sandra; Schmucker, Alexandra; Wolter, Stefanie; Zimmermann, Florian (2025): The Mannheim Enterprise Panel linked to the Establishment History Panel of the IAB 2010–2023 (MUP-BHP 1023). FDZ-Datenreport, (en), Nürnberg, DOI:.

Doherr, Thorsten (2023): The SearchEngine: A Holostic Approach to Matching. ZEW Discussion Paper No. 23-001, Mannheim.

Ganzer, Andreas; Schmucker, Alexandra; Stegmaier, Jens; Wolter, Stefanie (2023): Establishment History Panel 1975-2022. FDZ-Datenreport, 15/2023 (en), Nürnberg, 119 S. DOI:10.5164/IAB.FDZD.2315.en.v1.

Imprint

FDZ-Methodenreport 06|2025 EN

Date of publication
13 October 2025

Publisher
Research Data Centre (FDZ)
of the Federal Employment Agency (BA)
in the Institute for Employment Research (IAB)
Regensburger Str. 104
D-90478 Nuremberg

Download
https://doku.iab.de/fdz/reporte/2025/MR_06-25_EN.pdf

DOI
10.5164/IAB.FDZM.2506.en.v1

Dataset version
MUP-BHP 1023 v1, DOI: 10.5164/IAB.MUP-BHP1023.de.en.v1

All publications in the series "FDZ-Methodenreport" can be downloaded from
https://fdz.iab.de/en/research/publications/fdz-methodenreport-series/

Website
https://fdz.iab.de

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