

Discussion Paper No. 14-104

**The Mannheim Enterprise Panel (MUP)
and Firm Statistics for Germany**

Johannes Bersch, Sandra Gottschalk,
Bettina Müller, and Michaela Niefert

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 14-104

The Mannheim Enterprise Panel (MUP) and Firm Statistics for Germany

Johannes Bersch, Sandra Gottschalk,
Bettina Müller, and Michaela Niefert

Download this ZEW Discussion Paper from our ftp server:

<http://ftp.zew.de/pub/zew-docs/dp/dp14104.pdf>

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von
neueren Forschungsarbeiten des ZEW. Die Beiträge liegen in alleiniger Verantwortung
der Autoren und stellen nicht notwendigerweise die Meinung des ZEW dar.

Discussion Papers are intended to make results of ZEW research promptly available to other
economists in order to encourage discussion and suggestions for revisions. The authors are solely
responsible for the contents which do not necessarily represent the opinion of the ZEW.

The Mannheim Enterprise Panel (MUP) and Firm Statistics for Germany

Johannes Bersch, Sandra Gottschalk, Bettina Müller, and Michaela Niefert

Abstract

The Mannheim Enterprise Panel (*Mannheimer Unternehmenspanel* – MUP) of the Centre for European Economic Research (ZEW) is the most comprehensive micro database of companies in Germany outside the official business register (which is not accessible to the public). The MUP is based on the firm data pool of Creditreform e.V., which is the largest credit rating agency in Germany. At the end of 2013, the MUP contained information on 7.7 Mio firms, of which about 3.2 Mio were still active. Comparisons of the active stock of firms in the MUP with the Business Register of the Federal Statistical Office indicate that the MUP gives by and large a representative picture of the corporate landscape in Germany. The MUP is a valuable database for analyzing the number of start-ups and firm closures on a yearly basis for Germany. Further, the MUP is the sampling frame for the ZEW firm surveys and it is used for analyzing the development of firms over time.

Keywords: firm data, start-ups, firm closures, stock of firms in Germany

JEL Classification: C8, L26

1 Introduction

The Mannheim Enterprise Panel (**Mannheimer Unternehmenspanel** – MUP) of the Centre for European Economic Research (ZEW) is the most comprehensive micro database of companies in Germany. Since its foundation, ZEW co-operates with Creditreform e.V.¹, the largest credit rating agency in Germany, in order to build up and maintain the MUP. Since 2000, Creditreform's entire database is transferred to ZEW twice a year. Before that, from 1991 to 1999, Creditreform provided micro data on start-ups on a yearly basis. At ZEW, the data are cleaned, prepared for economic research and brought into a panel structure.

In this article, we describe how Creditreform data is made available for empirical research and how ZEW derives yearly start-up and firm closure numbers for Germany.

Creditreform data, and hence the MUP, contain the following information on companies headquartered in Germany including already closed companies: The complete address and telephone number, number of employed persons, amount of sales, legal form, five-digit industry sector code (NACE rev. 2), date of foundation, date of closure, shareholder structure and personal details about the involved persons (year of birth, gender, private address, profession, formal qualification, marital status, number of children) and Creditreform's credit rating score, and for a subset of medium sized corporations and company groups, balance sheet figures. ZEW adds county and municipal indicators, geocodes, and firm specific patent information. The MUP is also linked with the PROF²-database which is a database of companies which received government support for R&D projects. Furthermore, the MUP contains detailed firm specific data of insolvency procedures.

The statistical unit of the MUP is the legally independent enterprise. New businesses are registered by Creditreform when they have been

1. recorded in official registers (commercial register, register of associations, land register),
2. mentioned by print or internet media, business reports or
3. recognized by Creditreform in the course of investigations, which are typically initiated by requests of clients.

With this the MUP covers the population of German companies with "sufficient economic activity" to be noticed and registered by Creditreform. Companies with minor economic activities – such as freelancers, microenterprises, especially in the agricultural sector, and sideline businesses – are underrepresented in the MUP³.

The MUP is used for two purposes at ZEW. Firstly, it is used to calculate the yearly number of start-ups and business closures in Germany broken down by sector and region. The panel structure of the MUP also allows analyzing the development of the number of start-ups, business closures and the stock of companies over time (Müller et al., 2014). Secondly, ZEW uses

¹ The full German name is „Verband der Vereine Creditreform e.V.“

² Projektförder-Informationssystem

³ Egelin et al. (2012) showed that about 30% of start-ups which are registered by the MUP are run as a sideline business by the founder. Whereas the firm data bank of the Federal Statistical Office reports a rate of 40% among legally distinct start-ups.

the MUP as a sampling frame for several firm surveys, among others the Mannheim Innovation Panel (MIP, which is the German part of the Community Innovation Survey (CIS)), the Mannheim Start-up Panel, the Economic Survey of the Information Sector. Thus, ZEW survey data can be enriched by information from the MUP. In advance, the several half-year cross section data are joined to a panel which enables ZEW researchers analyzing the development of firms over time (latest, Müller et al., 2014).

This article is structured as follows. Section 2 describes the data processing steps which are necessary to use Creditreform data for the purposes of ZEW. In sections 3 and 4, we illustrate the methods we developed to calculate yearly numbers of start-ups and business closures for Germany. Finally, section 5 compares the MUP to the Business Register of the Federal Statistical Office in terms of the number of firms and the distribution of firms by size, legal form and economic activity.

2 Data Processing Steps

In order to use Creditreform data for the calculation of start-up and business closure numbers, as a sampling frame for firm surveys and for panel data analyses, several data processing steps are necessary:

1. The data have to be cleared up from invalid records. In most cases, these records are already marked by Creditreform as non-existing businesses or as dependent business units. A text analysis program developed by ZEW ('Texan') is used to detect further invalid entries. As a result of this step, nearly 2.9 out of 11 Mio records are deleted.
2. With the help of the software tool 'SearchEngine' which has been developed by ZEW (Doherr, 2014, see below) multiple entries of the same firm are identified. Roughly 4% of the companies which remained after the first clearing step are recorded more than once. This is due to the fact that Creditreform is decentralized in several independent units (2014: 133) which are each responsible for certain German regions (see for more details Engel and Fryges, 2002; Almus et al., 2000a, 2000b). Especially in border areas of these Creditreform districts, new firms are unintentionally registered by two or more different Creditreform units at the same time. Although new companies entries are reported to Creditreform's headquarter at a daily basis, multiple records are not apparent to the consultants in each case. The reasons can be different forms of spelling of the company name or company address. Furthermore, it is often the case that for multiple entries the private address of the owner is recorded instead of the company address. In addition, the relocation of small firms – in particular in other Creditreform districts – is sometimes not apparent to Creditreform and is therefore incorrectly registered as a start-up. After identification of invalid and multiple records these cases are deleted. At the end of 2013, this applied to almost 3.4 Mio entries so that the database included 7.7 Mio valid records at the end of data processing step 2.
3. The database of companies left after the clearing procedure includes active and inactive firms. The identification of the active companies requires several data processing steps which are described in detail in Section 4. At the end of 2013, these procedures led to the result that there were about 3.2 Mio active enterprises in Germany at that time.

In addition to these data clearing and data preparation steps, several procedures have been developed to improve the research potential of the MUP:

SearchEngine

The ZEW-Search-Engine is a text analysis software which has been programmed at ZEW (Doherr, 2014) and which is unique for the German empirical economic research community. This software is used to detect multiple records of the same firm and to combine micro data from different sources.

In general, no common identifier is available for combining different firm or individual data but only name and address information. These are prone to varying spellings which make a direct linkage difficult. The algorithm of the ZEW-Search-Engine matches dataset records via characteristic keywords in the name and address field. A rarely used word has a higher identification power than a frequently used one. Rarely used words are therefore defined as characteristic keywords. E.g.: "ZEW" is a characteristic keyword of the firm name "ZEW GmbH", whereas "GmbH" is not a characteristic keyword because it denotes a common legal form in Germany. The ZEW-Search-Engine detects all characteristic keywords which are then used to link data records.

Encoding of text information

The software tool Texan, which has also been developed by ZEW, is used to extract information from description fields which are attached by Creditreform to each firm record. Information in the description fields is not encoded and hence not directly available for data analyses. As an example, in 2013 start-ups has been identified which are producers of technologies or providers of services relevant for clean energy production or energy savings (Bersch et al., 2014). Young biotechnology firms have been identified in the same way (Rammer et al., 2006b).

ZEW Bankpanel

The MUP enables ZEW researchers to generate aggregate datasets in a variety of ways. One important example of such an aggregation is the ZEW Bankpanel. The MUP includes information for up to six banking relationships of a company, where the first relationship is denoted as the main bank ('Hausbank'), i.e. the bank used for day-to-day transactions. Creditreform determines the bank relationships by assigning each company to bank branches which are identified by a running number. The bank branches themselves are linked to the overall bank division by the unique German bank identifier BLZ. Using this link, ZEW constructs a panel of all banks operating in Germany. In order to adequately capture the banking market structure in Germany, mergers, acquisitions and bank market exits are investigated. Apart from the information on which banks are active in Germany and which banks merge, the ZEW Bankpanel can be enriched with other firm level data from the MUP. Thereby, ZEW makes use of the fact that the MUP covers almost the whole population of German companies. By aggregating all companies that have their main bank relationship with a certain bank, we are able to infer bank-specific information such as bank size, market share or portfolio shares by region and/or industry that goes way beyond balance sheet information usually provided for banks. ZEW is therefore able to give a clear picture of the structure of the corporate banking sector in Germany.

External information

On the basis of the industry sector code according to NACE rev. 2, which is assigned to each company by Creditreform, ZEW groups the companies into different technology areas. This grouping is based on definitions of high technology and knowledge intensive industry sectors in Germany (Gehrke et al., 2013, see also Table 2 in the Annex).

ZEW also enriches the MUP with several external publicly available or official data: county codes, municipal codes, geocodes, patent information and information from a database on firms which received government support for R&D projects (PROFI). The regional information is assigned via the address information of the company: postcode, town, and street. Geocodes are calculated by using the software ArcGIS which performs spatial analyses: On the basis of the company addresses ZEW computes longitudes and latitudes, distances between units and firm densities in specific regions. Patent information is made available by PATSTAT, a worldwide patent statistical database. The ZEW-Search-Engine is used to extract firm specific patent information from PATSTAT via a match of the names and addresses of the companies to the names and addresses of the patent applicants (see Aschhoff et al., 2013). The data base PROFIL, which lists government-funded companies is linked via ZEW-Search-Engine to the firm data of the MUP and via the MUP to ZEW survey data (especially to the firm data of the MIP). The data is prominently used as part of several ZEW research projects which evaluate the effects of public R&D subsidies. The linked data pool enables various evaluation studies (see e.g. Rammer et al., 2006a, Czarnitzki et al., 2003).

3 Number of Start-ups

Among others, the MUP is one of the major and most reliable sources for the number of start-ups in Germany. Most notably, the MUP covers information about the foundation date of start-ups. However, the information about the number of start-ups is not readily available when the data from Creditreform arrive at ZEW. The main reason is that Creditreform must have knowledge about the existence of newly founded firms. This can take some time so that start-ups are captured with a certain lag in Creditreform data. This leads to an underrepresentation of very young firms in the MUP in the most recent years. In order to overcome this underrepresentation, ZEW employs an extrapolation procedure.

The rate at which Creditreform gets information about the existence of a new firm depends mainly on the legal form. In Germany, there are two groups of legal forms, one for which a registration in the commercial register is obligatory and one for which the registration is voluntarily. The first group includes all corporations, i.e. limited liability companies (GmbH) and stock companies (AG), whereas the second group covers all other legal forms, such as business partnerships, and freelancers. If the founders of a new business either choose a legal form which requires a registration in the commercial register or if they register their business voluntarily then Creditreform gets notice of the new business within a relatively short period of time of less than one year. In all other cases it can take up to five years before Creditreform notices the new company. The reason is that Creditreform must actively search for these firms, which involves analyzing publications in the media or business reports or the following up of inquiries from

customers. As this second group involves a considerable number of new businesses (usually about 70% of each year's start-ups) an adjustment procedure for the latest number of start-ups is necessary in order to timely publish the number of new companies with the MUP. In general, registration lags have decreased over the last years, mainly due to the introduction of the electronic commercial registers⁴. These developments generate the necessity to adjust the mechanism such that it reacts to Creditreform's investigation speed.

Endogenous base date

In order to adjust for the differences in the registration lags ZEW employs an algorithm that builds on a variety of firm, region and industry specific factors in order to adequately extrapolate the number of start-ups. Like for any extrapolation procedure also for this algorithm a base year has to be determined. For the MUP, ZEW does not use a base year but rather a base date which consists of a year and a month. In the subsection "Extrapolation in more detail" we further explain how the base date is constructed and how extrapolation factors are applied. The base date t_b for the calculation of the number of start-ups with the MUP is defined such that each firm founded in t_b is assumed to have been captured up to t and serves as a rule of what fraction to extrapolate of the firms founded in $t, t-1, t-2, \dots, t_b + 1$. It determines the moment where the extrapolation starts. Only the number of start-ups in the years and month following the base date is extrapolated.

The search for a base date is crucial for any extrapolation mechanism. Frequently, base dates are arbitrarily set. In contrast to this, the base date in the extrapolation procedure for the calculation of the number of start-ups from the MUP is endogenous. In the extrapolation routine a rolling algorithm is applied that takes both the speed of search of Creditreform and the probability that a newly founded firm is investigated in some year $t_b + x$ into account. The result is a continuously changing extrapolation rule that takes the best available data into account.

The calculations are performed every time a new wave of Creditreform data has been delivered to ZEW. For each year of interest t , the full history of data that is available up to today is taken into account in the calculations. The base date for the most recent year of calculation and for each industry is calculated by applying a rule such that for any base date it must hold that the fraction of new firms founded at date t_b that are investigated in $t_b + x$ falls below a certain threshold value. As noted above, this rule then serves to determinate what fraction to extrapolate of the firms founded in $t, t-1, t-2, \dots, t_b + 1$.

Identifying Spin-Outs

The aim of the calculation of start-ups is to find an appropriate number of newly founded companies that are a risky investment for the entrepreneur(s). Corporate spin-outs that are founded by other enterprises and (though legally independent) are not fully exposed to the risk associated to start-ups are not in the focus of the analysis. ZEW applies the rule that all new firms that have more than 50 employees within the first two years are spin-outs. These firms are marked and not included in the extrapolation routine for the number of start-ups.

⁴ <https://www.handelsregister.de> or www.bundesanzeiger.de of the German "Bundesanzeiger" of the Ministry of Justice and Consumer Protection.

Treatment of missing values

Beside the necessity of extrapolating the number of start-ups in order to yield reasonable numbers, the data has to be cleaned up of missing values in several respects. Any kind of firm-specific information can be missing such as the firm's industry, its region but also its year of foundation. Most importantly, the year of foundation has to be imputed when it is missing because otherwise the number of start-ups will be downward biased. The algorithm uses a stratified bootstrap mechanism in order to assign a company to an industry, region or year of foundation. The stratum includes the date of registration by Creditreform, the firm's legal form as well as the firm's district⁵. The allocation of missing values is then carried out by drawing from the distribution of firms within the same strata. Bootstrapping and averaging thereby ensures that the mechanism is robust to outliers. The justification for such a mechanism is the observation that firms entering the dataset with missing year of foundation are usually assigned a year of foundation later on by Creditreform. If one compares the actually realized distribution of foundation years within a stratum to the one imputed, the aggregate distributions match very well.

Extrapolation in more detail

The extrapolation mechanism itself is based on the fact whether or not a firm must be registered in the commercial register (based on the legal form as explained above) and on its industry (see Table 2 for the applied industry classification at ZEW), i.e. for every point in time t ⁶ the algorithm calculates an extrapolation factor for each firm based on its industry and whether it is mandatorily registered. The extrapolation factor is at least one (and may be very high for firms at the youngest possible foundation month). Choosing industry as a stratum for extrapolation accommodates that industries may highly differ in their visibility (e.g. customer contact). After the extrapolation factors have been calculated, they are applied to the full dataset of actually observed start-ups. The number of start-ups is then calculated as the weighted (by extrapolation factor) sum of the firms in the data set.

Verification of extrapolated outcomes

The accuracy of the algorithm to determine the start-ups of a particular year t improves with each new wave of data because the probability that a newly founded firm is not covered decreases over time. Therefore, later waves can be used to evaluate the quality of the extrapolation algorithm. In the past, extrapolating year t once in t and again in $t+1$ usually shows less than 2% deviation compared to the maximum uncertainty outcome of year t .

Publication of results from the MUP

ZEW and Creditreform publish the number of start-ups in the biannual report 'Junge Unternehmen' ('Young Enterprises') (ZEW and Creditreform, 2014). Additionally, detailed

⁵ If regional data is not available on the level of the firm's district, the next higher regional aggregate is taken.

⁶ The extrapolation is not conducted on a yearly basis but rather makes use of a definition of periods that takes into account that the majority of firms is captured within one year after foundation and after two or three years only a small fraction of firms is added to a particular foundation year t . This calls for detailed time spans in the beginning and longer time spans for greater time distances both ensuring to capture the large dynamics in the early months after foundation and also to guarantee that for greater distances, the strata has sufficient size to deliver robust extrapolation.

indicators and graphics on regional and industry specific start-up activity can be acquired against a fee from ZEW.

4 Firm Closures and Stock of Firms

Just like the start-up numbers the numbers of firm closures cannot be directly taken from Creditreform data. The main reason is that not all companies are registered in the business register so that firm closures cannot be detected by simply looking at the de-registrations from the business register. To identify firm closures specific indicators generated by Creditreform are used. These indicators give information about the existence of the firms. One indicator covers the information whether a firm is deleted from the business register. Other closure indicators are the result of the investigations of Creditreform's consultants.

Bankruptcy and voluntary closure

Firms can leave the market in two different ways, bankruptcy and voluntary closure. Bankruptcy occurs when the firm is insolvent, while voluntary closure can have several reasons⁷. As in many other countries bankruptcy triggers an official act in Germany which are reflected in the business registers. Noticing bankruptcies therefore is rather easy for Creditreform. Creditreform gets information about insolvency procedures from an official internet platform⁸ where announcements of insolvency courts are published. This results in the fact that the MUP covers information about the insolvency procedure for each affected firm. These include: date of application, initiation, sequestration, cessation, suspension of the procedure and information on potential insolvency plan procedures. Firms which are marked by Creditreform of being in an insolvency procedure are identified as closed in the MUP. The date of closure is the date of the insolvency application.

Voluntary closed firms in the MUP are those firms which are closed according to Creditreform's closure indicators and are not insolvent. The MUP closure date for these firms corresponds to the date Creditreform's closure indicator emerges the first time in the data base taking into account the registration lag in the Creditreform data (see Section 3 and Engel and Fryges, 2002). Considering the registration lag is relevant for those firms that are not obliged to register in the commercial register (freelancers, microenterprises, sole traders, business partnerships). For these firms we assume a registration lag of one year. The closure date is then the year of the first appearance of Creditreform's closure indicator minus one. For corporations (GmbH, UG, AG) and business partnerships which must register (OHG, KG, GmbH & Co. KG) and companionships we assume no registration lag.

As we receive the total Creditreform database only since 1999, we are not able to observe voluntary closures before that date. Further, a comparison of the time series dynamic of voluntary closures on the one hand with the time series dynamic of business de-registrations in official statistical data on the other reveals an under coverage of firms in the MUP until 2002.

⁷ The main reason for closing a venture voluntarily is financial distress (Egeln et al, 2010). Other reasons are too small profits and family or health related reasons.

⁸ www.Insolvenzbekanntmachungen.de

We are able to fully recognize the stock of firms since 2002 and the number of voluntary closures for the year 2003 onwards.

Potentially closed firms

Just like in the case of start-ups the registration lag in the Creditreform data leads to a potential under-reporting of firm closures in the most recent years. Again, the under-reporting is related to the group of firms which are not obliged to register in the commercial register. In addition, we observe that records of very small firms (microenterprises and freelancers) are updated with very low frequency. A high number of these records have never been updated since the first entry. Manual investigations with random samples showed that the probability is high that the firms behind these records do no longer exist. This has to be taken into account when calculating the number of firm closures with the MUP.

As there is no other information available than the updating frequency the following assumption is made with respect to closure of non-registered and very small firms:

Firms that are not obliged to register in the commercial register are regarded as closed when their records not been updated within the last three years. The records of these firms are marked with “potentially closed”. The closure date for these firms is set to the date of the last update plus one year.

To account for the uncertainty of this assumption, an uncertainty parameter is calculated for each potentially closed firm. This uncertainty parameter says how likely it is that a record of a firm keeps the potentially-closed status. It is calculated by analyzing the development of potentially closed firms since 1999. On average, 70% of the potentially closed firms keep the status of voluntary closure each year; for 10% it becomes apparent that they are indeed closed and they get retrospectively marked by the closure indicator. Another 8% are detected as incorrect records, i.e. records that have a firm identifier but which do not belong to a firm. The remaining 12% of potentially closed firms turn out to be economically active because Creditreform updated information for these firms in the respective year.

The probability of keeping the status ‘potentially closed’ decreases over time. Firms that have been in this status for ten years keep it with a probability of 80%. If a firm gets this status only recently it keeps it with a probability of 50%. The uncertainty parameter therefore varies between 50 and 80%.

Because we require that the records are not updated within the last three years for identifying the potentially closed companies we cannot calculate the number of firm closures for the most recent years with the procedure described above. In order to being able to report timely numbers of firm closures the numbers of potentially closed firms are projected according to the growth rates of the voluntary closed firms between the current year and the year before. The growth rates are differentiated by industry sector groups according to Table 3. On average, 80% of firm closures are voluntarily, 20% are bankruptcies. From the voluntary closed firms about 40% are marked as potentially closed. These percentages do not vary much over time.

Stock of firms

Since 2000, ZEW gets the total stock of firms recorded in the datasets of Creditreform. The stock of German firms can be firstly reported for the year 2002. The time series is computed in adding

the difference between the number of start-ups plus spin-outs and the number of closures for each year, respectively.

Table 1 summarizes the development of the numbers of start-ups, firm closures and active firms over time. The numbers refer to firms of the industries listed in Table 3 (i.e. firms from the agricultural and public sector are not included). The numbers of genuine start-ups are shown in column two. In column three the numbers of spin-outs - new firms which are outsourced units from existing companies - are depicted. Column four shows the total numbers of closed firms. Starting from 2002 with a total of 3,157,000 German firms, column five lists the resulting stock of firms per year. Until 2008, the stock of firm increased, after that the number of active companies in Germany declined steadily up to a number of 3,230,000 companies at the end of 2013.

Table 1 Firm Dynamics in Germany 2002-2013

Year	Start-ups	Spin-outs	Closures	Stock of firm
2002	228,000	1,700		3,156,000
2003	249,000	1,700	200,000	3,207,000
2004	273,000	1,600	211,000	3,271,000
2005	248,000	1,500	206,000	3,315,000
2006	236,000	1,500	224,000	3,329,000
2007	217,000	1,400	201,000	3,346,000
2008	203,000	1,300	193,000	3,357,000
2009	216,000	1,200	222,000	3,352,000
2010	208,000	1,100	222,000	3,340,000
2011	195,000	900	226,000	3,309,000
2012*	170,000	800	194,000 ¹⁾	3,286,000 ²⁾
2013*	155,000	600	213,000 ¹⁾	3,230,000 ²⁾

Source: Mannheim Enterprise Panel (MUP), ZEW Mannheim

* For 2012 and 2013, the numbers of closures and stock of firms are estimated and therefore preliminary.

1) potential overestimation of the number of closures, 2) potential underestimation of the stock of active firms.

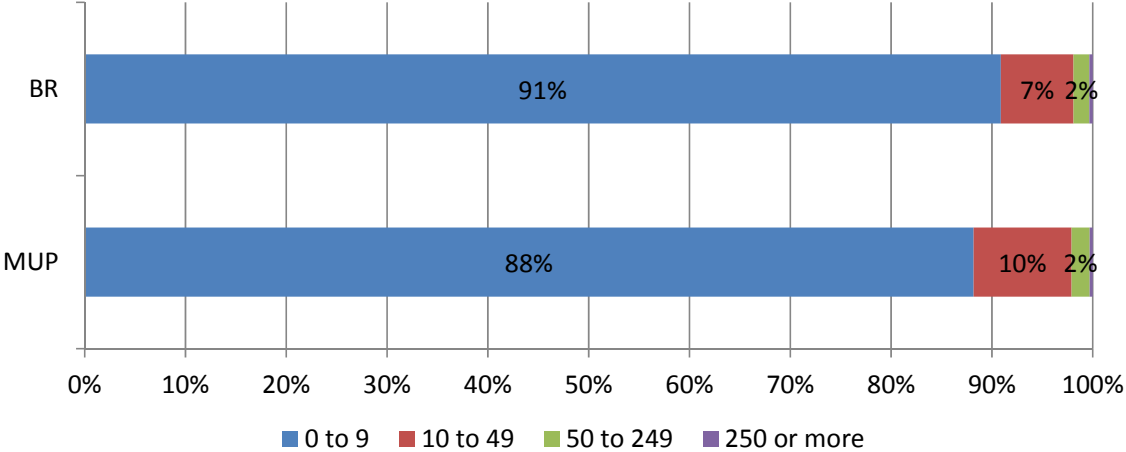
5 Comparison of the MUP with the Business Register

In order to assess the extent to which the MUP covers the total stock of firms in Germany, it is compared with the Statistical Business Register of the Federal Statistical Office. The Business Register (BR) is primarily based on data of the fiscal authority and the Federal Employment Agency. It captures firms which are either subjected to VAT or have at least one employee subject to social insurance contribution. Like in the MUP, the statistical unit is the legally independent company (Rink et al. 2013). By the end of 2012, the BR reports a stock of 3,663,000

firms, compared to 3,286,000 firms included in the MUP.⁹ Thus, the MUP contains about 377,000 firms less than the BR. Assuming that the BR represents the full stock of German firms, the MUP captures 90% of it.

Comparing the distribution of firms in the MUP and the BR by size, legal form and economic activity indicates in which parts the largest differences occur. Overall, the distribution by firm size based on employment looks quite similar for the two data sets. The lion’s share of firms has an employment figure of less than 10. This share is somewhat lower in the MUP (88%) than in the BR (91%), while the proportion of firms with an employment of 10 to 49 is higher in the MUP (10% versus 7% in the BR). This higher share in the MUP can partly be attributed to the fact that the employment figure in the MUP refers to persons employed (i.e. including self-employed persons such as the owners of the firm) whereas employment in the BR refers to employees only. The share of firms falling into the upper size classes are almost the same in the two data sources (Figure 1).

Figure 1 Distribution of Firms by Firm Size

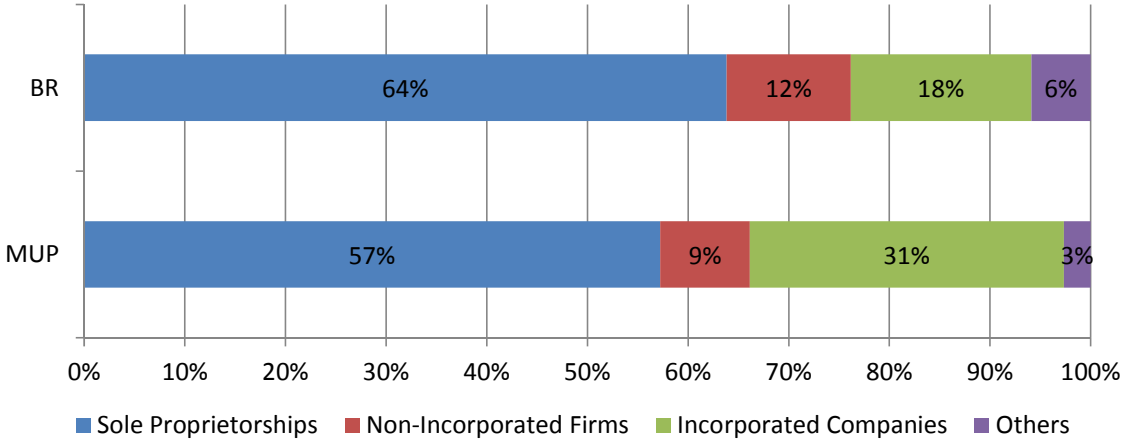


Source: Mannheim Enterprise Panel (MUP), ZEW; Statistical Business Register, Federal Statistical Office

The fact that the firms in the MUP are relatively weakly represented in the smallest size class is also reflected in the distribution by legal form. Only 57% of the MUP firms are sole proprietorships compared to 64% of the BR firms. Sole proprietorships fall practically completely into the class of firms with less than 10 employees. The MUP also shows a lower share of non-incorporated firms. In contrast, the proportion of incorporated companies - the legal form typically chosen by larger firms - is considerably higher in the MUP (31%) than in the BR (18%) (Figure 2).

⁹ Certain economic activities are not captured by the BR. These are agriculture, forestry and fishing, public administration and defense, compulsory social security, activities of private households and activities of extraterritorial organizations and bodies. The MUP does not (fully) capture these activities, either. They are excluded from the analysis and are not contained within the stock of firms as specified above.

Figure 2 Distribution of Firms by Legal Form



Source: Mannheim Enterprise Panel (MUP), ZEW; Statistical Business Register, Federal Statistical Office

The distribution of firms by economic activity looks relatively similar for the MUP and the BR. The most striking differences are that the MUP exhibits somewhat higher shares in construction, trade, and financial and insurance activities, whereas it shows lower shares in accommodation and food service activities, real estate activities, and consumer-related services (education, human health and social activities, arts, entertainment and recreation, and other service activities) (Table 2). The latter activities are particularly often conducted by very small firms which are not obliged to sign in the commercial register and about which there is hardly any public information available. These micro firms are, for example, smaller cafés, snack bars, freelance real-estate agents, consultants, providers of residential care services, own-account journalists, and beauty salons.

Table 2 Distribution of Firms by Economic Activity

Economic Activity	MUP	BR
Mining and Quarrying	0.07%	0.06%
Manufacturing	7.18%	6.90%
Electricity, Gas, Steam and Air Conditioning Supply	0.73%	1.65%
Water Supply; Sewerage, Waste Management and Remediation Activities	0.32%	0.34%
Construction	13.33%	10.72%
Wholesale and Retail Trade; Repair of Motor Vehicles and Motorcycles	21.73%	18.30%
Transportation and Storage	3.82%	3.33%
Accommodation and Food Service Activities	5.21%	6.79%
Information and Communication	3.53%	3.57%
Financial and Insurance Activities	4.29%	1.91%
Real Estate Activities	4.57%	8.86%
Professional, Scientific and Technical Activities	14.45%	14.06%
Administrative and Support Service Activities	6.53%	5.55%
Education	1.30%	2.09%
Human Health and Social Work Activities	4.92%	6.49%
Arts, Entertainment and Recreation	2.13%	2.86%
Other Service Activities	5.90%	6.51%
Total	100.00%	100.00%

Source: Mannheim Enterprise Panel (MUP), ZEW; Statistical Business Register, Federal Statistical Office

All in all, comparing the distribution of firms by size, legal form and economic activity indicates that the difference in the absolute number of firms in the MUP and the BR can be explained by the fact that the MUP captures micro firms to a disproportionately small extent. These firms are included in the BR as soon as their sales volume comes up to 17,500 Euro per year so that they are subjected to VAT. However, they are only gathered in the MUP if they are either recorded in official registers, if there is public information available about them, or if they are recognized by Creditreform via customer requests. This will often not be the case even if they are achieving sales of 17,500 Euros. Thus, the difference in the absolute number of firms between the MUP and the BR is primarily due to the different information sources and therefore different selection criteria used by Creditreform and the Federal Statistical Office. But it can be stated that the distributions of firms by the mentioned indicators in the MUP and the BR are overall rather similar despite the somewhat smaller stock of firms in the MUP. By and large, the MUP gives a representative picture of the corporate landscape in Germany.

6 Literature

- Almus, M., D. Engel, and S. Prantl (2000a), The „Mannheim Foundation Panels“ of the Centre for European Economic Research (ZEW), ZEW Documentation 00-02, Mannheim.
- Almus, M., D. Engel, and S. Prantl (2000b), The ZEW Foundation Panels and the Mannheim Enterprise Panel (MUP) of the Centre for European Economic Research (ZEW), Schmollers Jahrbuch 120, 301-308.
- Aschhoff, B., G. Licht, and P. Schliessler (2013), Who drives smart growth? The contribution of small and young firms to inventions in sustainable technologies, WWWforEurope Project Working Paper No 47, Brussel.
- Bersch, J., J. Egel, D. Faustmann, D. Höwer, B. Müller, and M. Murmann (2014), Potenziale und Hemmnisse von Unternehmensgründungen im Vollzug der Energiewende, Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie, Mannheim.
- Czarnitzki, D., T. Doherr, A. Fier, G. Licht, and C. Rammer (2003), Öffentliche Förderung der Innovationsaktivitäten von Unternehmen in Deutschland, Studien zum deutschen Innovationssystem, 17-03, Mannheim.
- Doherr, T. (2014), ZEW-Search-Engine, ZEW Documentation, Mannheim, *forthcoming*.
- Engel, D. and H. Fryges (2001), Aufbereitung und Angebot der ZEW Gründungsindikatoren, ZEW Documentation 02-01, Mannheim.
- Egel, J., U. Falk, D. Heger, D. Höwer, and G. Metzger (2010), Ursachen für das Scheitern junger Unternehmen in den ersten fünf Jahren ihres Bestehens, Studie im Auftrag des Bundesministeriums für Wirtschaft und Technologie, ZEW, ZIS Universität Mannheim, Creditreform, Mannheim, Neuss.
- Egel, J., S. Gottschalk, D. Höwer, and B. Müller (2012), Existenzgründungsgeschehen in Bayern, Report on behalf of Bayerisches Staatsministerium für Wirtschaft, Infrastruktur, Verkehr, Technologie, Munich.
- Gehrke, B., R. Frietsch, C. Rammer, and M. Leidmann (2013), Neuabgrenzung forschungsintensiver Industrien und Güter, NIW/ISI/ZEW Listen 2012, Hrsg.: Expertenkommission Forschung und Innovation (EFI), Studien zum deutschen Innovationssystem 8-2013, Berlin.
- Müller, B., S. Gottschalk, M. Niefert, and C. Rammer (2014), Unternehmensdynamik in der Wissenswirtschaft in Deutschland 2012, Studien zum deutschen Innovationssystem, 4-2014, Mannheim.

Rammer, C., V. Zimmermann, E. Müller, D. Heger, B. Aschhoff, and F. Reize (2006a), Innovationspotenziale von kleinen und Mittleren Unternehmen, ZEW Wirtschaftsanalysen, 79, Baden-Baden.

Rammer, C., J. Ohmstedt, H. Binz, and O. Heneric (2006b), Unternehmensgründungen in der Biotechnologie in Deutschland 1991 bis 2004, ZEW Dokumentation 06-03, Mannheim.

Rink, A., Seiwert, I., and R. Opfermann (2013), Unternehmensdemografie: methodischer Ansatz und Ergebnisse 2005 bis 2010, Wirtschaft und Statistik, Juni 2013, 422-439.

ZEW and Creditreform (2014), Junge Unternehmen, 1, Neuss und Mannheim.

7 Annex

Table 3: Classification of industry sector groups

	Industry Sector Groups	Industry sector classification (NACE rev. 2)
1	Cutting-edge technology manufacturing	20.2, 21, 24.46, 25.4, 26.11, 26.2, 26.3, 26.4, 26.51, 26.6, 26.7, 30.3, 30.4
2	High-technology manufacturing	20.13, 20.14, 20.16, 20.42, 20.51, 20.53, 20.59, 22.11, 23.19, 23.44, 26.12, 27.11, 27.12, 27.2, 27.31, 27.33, 27.4, 27.9, 28.11, 28.12, 28.13, 28.15, 28.23, 28.24, 28.29, 28.3, 28.41, 28.49, 28.92, 28.93, 28.94, 28.99, 29.1, 29.31, 29.32, 30.2, 33.2
3	Non-high-tech manufacturing	10-33 (ex sectors 1 and 2)
4	Technology-intensive services	61.1-61.3, 62, 63.1, 71.1, 71.2, 72.1
5	Non-technical consulting	69, 70.2, 72.2, 73
6	Other business-oriented	61-63, 69-72, 77.1, 77.3, 77.4, 78, 80, 81 (ex 70.1, 74.2)
7	Consumer-oriented services	55-56, 58-60, 68, 74.2, 75, 77.2, 79, 85.5-85.6, 86-88, 90-93, 95-96
8	Energy/Mining/Disposal	5-9, 35-39
9	Construction	41-43
10	Trade	49-52
11	Traffic/Mailing	49-53
12	Mailing	53
13	Banks/Insurances	64-66 (ex 64.2)

Source: Own classification, NIW/ISI/ZEW Listen 2012 (Gehrke et al., 2013)