Identification of Local Industrial Clusters in Germany

THOMAS BRENNER

Evolutionary Economics Unit, Max-Planck-Institute for Research into Economic Systems, Kahlaische Str. 10, D-07745 Jena, Germany. Email: Brenner@econ.mpg.de

BRENNER T. (2006) Identification of local industrial clusters in Germany, Regional Studies 40, 1–14. This paper presents a method that allows local industrial clusters to be identified and applies this method to Germany. The method is applied to all three-digit manufacturing industries in Germany. The results are used in two ways. First, they provide some information about which industries are clustering, and which are not. Second, a complete list of all local clusters that existed in Germany in 2001 and are identified by this method is given. The spatial distribution of these local clusters in Germany is discussed in the light of other studies on Germany.

Local industrial clusters Industry study Spatial agglomeration Empirical methodology

INTRODUCTION

Local clusters and industrial districts have been studied intensively in the recent economic literature. These studies mainly aim to identify the prerequisites for the development of such local systems and the specific characteristics responsible for their economic success. It is usually implicitly assumed that local clusters and
industrial districts can be clearly identified. This means that it is assumed that local systems can be classified into two categories according to their economic situation: successful regions, labelled local clusters or industrial districts, and regions lagging behind.

In the literature on case studies there is not usually much discussion about the identification of local clusters. Those who conduct the case studies implicitly, and usually rightly so, assume that they have correctly identified the local cluster they are studying. There are other approaches that try to identify all local clusters within a country (Sforzi, 1990; Isaksen, 1996; Paniccia, 1998; Braunerhjelm and Carlsson, 1999; Sternberg and Litzenberger, 2004). In these approaches a threshold level of the number of firms or employees in a region is arbitrarily defined, usually in relation to the size of the region. All regions in which the number of firms or employees in an industry exceeds this threshold value are said to contain a local cluster. In most of the approaches additional conditions are formulated. The most elaborated approach has been developed for Italy (Sforzi, 1990; Sforzi et al., 1997). However, even in this case the condition on the size of the industrial agglomeration is little discussed. Furthermore, the same condition is applied to all industries.

What is missing is a theoretical or empirical explanation for the threshold level used. There are several theoretical works that analyse what the spatial distribution of firms should look like if there is the phenomenon of clustering (e.g. Ellison and Glaeser, 1997, 1999; Bottazzi et al., 2002; Dumas et al., 2002; for a discussion of the different approaches, see Brenner, 2006). However, these studies only aim to show that firms are not randomly distributed in space and, therefore, agglomeration forces are present. They do not aim to identify the locations of agglomerations.

This paper provides a methodology to identify empirically the threshold in the number of firms that separates those regions containing a local cluster in a certain industry from those that do not. The methodology used here is based on a theoretical model of clustering (Brenner, 2001) and some methodological considerations (Brenner, 2006). It allows predictions to be made about the shape of the distribution of firms among regions. Through fitting the parameters of this distribution function, two results are obtained. First, it can be tested whether local clusters exist in a certain industry or whether the spatial distribution of firms can be explained without assuming clustering forces. This corresponds to the approaches on the spatial distribution of firms in the literature (Ellison and Glaeser, 1997). Second, the threshold for the number of firms in local clusters can be calculated based on the parameters of the distribution function. This means that the threshold is obtained empirically and for each industry separately. This value can then be used to identify existing local clusters.

The methodology is then applied to Germany. It is applied at the level of administrative districts and conducted for each three-digit manufacturing industry separately. The results are used in two ways. First, they answer the question about whether clustering exists in the different manufacturing industries in Germany. Second, the location of clusters in the different industries in Germany can be identified. This provides a picture of clustering in Germany. A similar study has been conducted for Germany by Sternberg and Litzenberger (2004). In contrast to the study conducted herein, they applied the usual approach using a cluster-index based on the share of an industry in a region. Furthermore, they used a higher aggregation on the industrial as well as the regional level.

The paper is structured as follows. In the second section the theoretical framework is outlined and predictions for the distribution of firms among regions are presented. The methodology to test these predictions is developed in the third section. In the fourth section this methodology is applied to Germany and the results are discussed. The fifth section concludes.

### THEORY AND PREDICTIONS

In order to identify whether industries show clustering in space, a model is needed that describes both a situation with clustering and a situation without clustering. This model can then be used to make two predictions about the spatial distribution of firms, one for the situation with clustering and one for the situation without clustering. Finally, the adequateness of the two predictions can be compared. There are several such models in the literature (Ellison and Glaeser, 1997; Brenner, 2001; Bottazzi et al., 2002; Dumas et al., 2002). Here the model by Brenner (2001) is used. It is beyond the scope of the present paper to describe the whole model or how the predictions are derived. Hence, only some intuitive arguments are given. For further details, see Brenner (2001, 2004, 2006).

#### Basic assumptions

The theoretical considerations used here to predict the firm distribution among regions is based on two fundamental assumptions. The first separates those industries in which local clusters exist from those industries without local clusters. It is assumed that industries with local clusters are characterized by local self-augmenting processes that, in some regions, lead to much higher economic activity, and therefore to a much higher number of firms than in other regions (for a theoretical analysis of this, see Brenner, 2001). Let us denote the probability that a randomly chosen region of size s contains f firms by P(f|s) and call it the ‘firm distribution among region’. Then, the above assumption leads to the predictions that P(f|s) is
bimodal. This means that there are many regions with a small number of firms but also quite a number of regions with a high number of firms. Between these cases there are numbers of firms that are rarely observed (for an example of such a distribution, see Fig. 1). Industries in which no local clusters exist are characterized by a unimodal firm distribution \( P(f|s) \).

The second assumption concerns the functional form of the firm distribution \( P(f|s) \). In the literature the functional form of the firm distribution is deduced from theoretical considerations about the stochastic process of firm location (e.g. Ellison and Glaeser, 1997; Bottazzi et al., 2002; Dumais et al., 2002). Here it is assumed that, in addition to the stochasticity of firm locations, the attractiveness of regions is influenced by many factors that cannot be explicitly considered in an approach such as the one used here. Examples are the human capital available in the region, cultural factors, and geographic attractiveness. The only factor that is explicitly considered, as in most other studies, is the size of the regions, \( s \). Therefore, the firm distribution is denoted by \( P(f|s) \), which means that for each size a different distribution is assumed. The other factors have to be treated as stochastic influences as well. Usually the distribution of these factors among regions is not even known. For some factors the functional form of the distribution can be at least estimated based on empirical data (Brenner, 2006). However, it is argued here that the final decision about the functional form of the firm distribution has to be made based on empirical data. Brenner (2006) tests three different functional forms: an exponentially decreasing function; a binomial distribution; and a Boltzmann distribution. The binomial and the Boltzmann distributions have quite similar shapes. Therefore, only one of these distributions is used here. The Boltzmann distribution has been found to describe the highest percentage of the empirical distribution (Brenner, 2006). Therefore, a mixture of an exponentially decreasing function and a Boltzmann distribution is used here for \( P(f|s) \).

Predictions for the industrial firm distribution among regions

Combining the two assumptions that have been outlined above, it is possible to formulate mathematical predictions for the industrial firm distribution among regions. First, the prediction for an industry without local clusters is formulated. In this case, the firm distribution should follow the distribution of the factors that influence the attractiveness of regions plus the stochastic fluctuations caused by the location decision of firms. It is argued above that it is impossible to deduce the functional form of the resulting distribution theoretically. Therefore, a mixture of an exponentially decreasing function and a Boltzmann distribution is used that has been found to describe the empirical situation quite well (Brenner, 2006). This means that the firm distribution among regions is predicted by:

\[
P_n(f|s) = (1 - \xi_1(s)) \cdot \xi_1(s)^f + \xi_2(s) \cdot f \cdot \xi_2(s)^f
\]

for industries with no local clusters. \( \xi_1(s), \xi_2(s) \) and \( \xi_3(s) \) are parameters that depend on the size of the considered region, \( s \). The first term on the right-hand side of equation (1) is an exponentially decreasing function. The second term is a Boltzmann distribution. Distribution (1) is called the natural distribution.

The second case in which local clusters occur is characterized by the fact that some regions might contain a much higher number of firms than the others. The theory predicts that there is a gap between the number of firms in those regions that contain an industrial cluster and those regions that do not contain such a cluster. Mathematically such a situation can be described by the following distribution:

\[
P_c(f|s) = (1 - \xi_1(s) - \xi_2(s)) \cdot \xi_1(s)^f + \xi_2(s) \cdot f \cdot \xi_2(s)^f + \xi_3(s) \cdot f \cdot \xi_3(s)^f
\]

\[
= \begin{cases} 
\xi_3(s) \cdot f \cdot \xi_3(s)^f - \xi_2(s) & \text{if } f > \xi_3(s) \\
\xi_2(s) \cdot f \cdot \xi_2(s)^f & \text{if } f \leq \xi_3(s)
\end{cases}
\]

where \( \xi_1(s), \xi_2(s), \xi_3(s) \) are the parameters of the function. This distribution is called the cluster distribution. It consists of the same two distributions as the natural distribution. In addition, a further Boltzmann-like distribution is added that takes values greater than zero only for high firm numbers \( f \). All industries that contain local clusters should be better described by the cluster distribution. This will be tested for all manufacturing industries and some service industries below.

However, testing the predictions for the distribution of firms does not allow statements to be made about the existence of local clusters. Brenner (2004) defines local industrial clusters as spatial agglomerations of the firms in an industry in which the firms benefit from their co-location. Here only the existence of spatial agglomerations of firms is checked. There might be other reasons for the existence of such agglomerations, such as natural resources or urban externalities.
However, different reasons for industrial agglomerations might cause the mathematical shape of the firm distribution to be different. Here the adequateness of the particular shape that is predicted by the above theory of local industrial clusters is tested. Still other reasons for industrial agglomerations might cause the same mathematical shape of the firm distribution, but this is at least less likely. Nevertheless, it cannot be excluded that the method used here identifies local agglomerations that are not caused by clustering forces but by other factors.

**Influence of the size of regions**

In the approach taken here the size of regions is considered explicitly. This implies that the firm distribution, $P(f|s)$, is different for regions of different sizes. However, some assumptions about the impact of the size on the firm distribution are necessary. If $P(f|s_1)$ and $P(f|s_2)$ are completely independent of each other, the empirical data would provide only one datum point for each firm distribution, $P(f|s)$. This would not allow any statements to be made.

The assumption used here is that the average expected number of firms in a region increases linearly with the size of the region. This is an assumption that is also used in other studies (e.g. Ellison and Glaeser, 1997). In contrast, the shape of the firm distribution is assumed to be independent of the size of the region. The shape of the firm distribution is determined by the functional form, which is defined above, and the share that each function contributes: $\xi_j(s)$ and $\xi_s(s)$. This implies that $\xi_j(s)$ and $\xi_s(s)$ are assumed to be independent of $s$:

$$\xi_j(s) = \xi_j \quad \text{and} \quad \xi_s(s) = \xi_s$$  (3)

The average expected number of firms according to each term in equations (1) and (2) is given by the parameters within the three functions, namely $\xi_1(s)$, $\xi_2(s)$, $\xi_4(s)$ and $\xi_5(s)$. If one calculates the average expected number of firms using only the first term on the right-hand side of equation (1) or (2), one obtains:

$$\frac{\xi_1(s)}{1 - \xi_1(s)}$$  (4)

One would like this average expected number of firms to increase linearly with the size $s$ of the region, meaning that one would like to obtain an average number of:

$$\xi_1 \cdot s$$  (5)

Equating (4) and (5), one obtains:

$$\xi_1(s) = \frac{\xi_1 \cdot s}{1 + \xi_1 \cdot s}$$  (6)

which means that if one defines parameter $\xi_1(s)$ in such a way, the first term in equations (1) and (2) produces an average value that increases linearly with $s$. The same will also be performed with the second term in these equations. The average expected number of firms according to the second term results is $(2 \cdot \xi_2(s))/(1 - \xi_2(s))$. With the same calculation as above, one obtains:

$$\xi_2(s) = \frac{\xi_2 \cdot s}{2 + \xi_2 \cdot s}$$  (7)

as the definition of $\xi_2(s)$ that guarantees that the second term’s average prediction increases linearly with the size of the region $s$. The average expected number of firms according to the last term in equation (2) is given by $\xi_4 + (2 \cdot \xi_5(s))/(1 - \xi_5(s))$. There are several ways to make this value linearly dependent on $s$. The most obvious is to assume that the first term, $\xi_4$, depends linearly on $s$:

$$\xi_4(s) = \xi_4 \cdot s$$  (8)

and that the second term also depends linearly on $s$, which is reached by:

$$\xi_5(s) = \frac{\xi_5 \cdot s}{2 + \xi_5 \cdot s}$$  (9)

As a consequence, if one uses these definitions for parameters $\xi_1(s)$, $\xi_2(s)$, $\xi_4(s)$, $\xi_5(s)$, $\xi_6(s)$, the average value of each of the terms in equations (1) and (2) depends linearly on the region’s size, so that the average predicted number of firms also depends linearly on $s$. Through this a new set of parameters is obtained, namely $\xi_1$, $\xi_2$, $\xi_4$, $\xi_5$, and $\xi_6$, which are independent of the size of regions and can be fitted to the empirical data.

**EMPIRICAL DATA AND METHOD OF ANALYSIS**

**Empirical data**

The data used in this approach were collected by the German Federal Institute for Labour. The data set contains the number of firms for each three-digit industry and each of the 441 administrative districts in Germany for 30 June 2001. The study conducted herein is restricted to the 104 manufacturing industries. The industries are denoted by $i \in \{1, 2, \ldots, 104\}$. Some of the empirical analysis is also conducted for a few service industries that are frequently discussed in the literature. The administrative districts are denoted by $r \in \{1, 2, \ldots, 441\}$. The number of firms in each industry and districts is denoted by $f(i, r)$.

In the approaches that identify local clusters in the literature, the number of firms or employment is
usually compared with the ‘natural’ share of the region (e.g. Sforzi, 1990; Isaksen, 1996; Paniccia, 1998; Braunerhjelm and Carlsson, 1999; Sternberg and Litzenberger, 2004). This implies that the number of firms has to be studied in relation to the size of the respective region. This seems to be plausible for the approaches taken in the literature. However, arguments can be put forward in favour of using the absolute number of firms. On the one hand, it is assumed here that the existence of local industrial clusters is caused by local self-augmenting processes. These result from the symbiotic interaction among firms and between firms and local circumstances. The effects of these interactions depend on the number of firms but not on the total size of regions in terms of population or employees. Therefore, the self-augmenting processes should occur whenever the absolute number of firms exceeds a certain value. On the other hand, the total number of employees in a region might well determine the number of firms that might be established in a region, once a cluster emerges. The relative number of firms takes this into account.

Therefore, the firm distribution among regions is studied below considering the size of regions as well as neglecting their size. To this end, the size of regions, \( r \), is defined in two ways. First, the size of a region is defined as the share of all employees in Germany that are employed within the region:

\[
s(r) = \frac{\sum c(i, r)}{\sum \sum c(i, r)} \tag{10}
\]

where \( c(i, r, t) \) is the number of employees in industry, \( i \), and region, \( r \), at time, \( t \). Second, the size of all regions is set to the same value, namely:

\[
s(r) = \frac{1}{441} \tag{11}
\]

This allows the same mathematical formulations to be used in the empirical study.

Restrictions on the models

The first two terms of the natural distribution (1) and the cluster distribution (2) are identical, except for the factor in front of the exponential part. The last term of the cluster distribution is defined in a similar way to the second term with the form of a Boltzmann-like distribution. It is a shifted Boltzmann-like distribution, which is zero for all numbers of firms below \( \xi_i \cdot s \) and has the shape of a Boltzmann-like distribution for all numbers of firms above \( \xi_i \cdot s \). Hence, this term represents the clusters that exist.

The parameters \( \xi_i \) and \( \xi_0 \) are restricted in the analysis below. The shifted Boltzmann-like distribution is designed to describe the local clusters existing in an industry. This implies several characteristics of this part of the distribution. First, it should only describe a few regions. Local clusters have to be the exception. One would not talk about ‘local industrial clusters’ if they occurred in most of the regions. Hence, the share of the regions that contain a cluster has to be small. A total of 10% is assumed to be the maximal share that is accepted. This implies that \( \xi_0 \leq 0.1 \) has to be satisfied.

Second, only those regions with a high number of firms should be classified as local industrial clusters. Nevertheless, the line between clusters and other regions should be empirically determined. It is represented by the parameter \( \xi_4 \). A value of \( \xi_4 = J \), with \( J \) denoting the total number of firms in the industry under consideration, would imply that all regions containing a number of firms above average would be called clusters. This seems to be inadequate. In the literature a values of \( \xi_3 = 3 \) or \( \xi_4 = 4 \) are used (Isaksen, 1996; Sternberg and Litzenberger, 2004). Here, a more complicated restriction is chosen for the parameter \( \xi_4 \).

According to the analysis in Brenner (2001), every region that includes a cluster of the industry under consideration should contain more firms than any region that does not contain such a cluster. This implies that the part of the distribution described by the last term in equation (2) should be separated from the rest of the distribution. In reality such a separation cannot be expected because regions differ with respect to factors other than their attractiveness. Nevertheless, the overlapping of the two parts of the distribution should be small. The amount of overlap is given by the number of regions that contain a number of firms higher than \( \xi_4 \cdot s \) according to the first and second part of the cluster distribution. This number is denoted by \( n_{\xi_4} \) here. It should be small compared with the number of regions that contain industrial clusters. Otherwise, most of the regions containing many firms could be explained without considering the phenomenon of clustering. Therefore, \( \xi_4 \) is restricted to those values for which at least five-sixths of the regions that contain a number of firms higher than \( \xi_4 \cdot s \) are explained by the cluster part of distribution (2). This is reached by the condition:

\[
0.2 \cdot \xi_0 > \frac{n_{\xi_4}}{N_r} \tag{12}
\]

Method to test the theoretical distributions

To fit the two theoretical distributions to reality, the likelihood value is calculated for each parameter set using the data on each industry. The likelihood value is the probability that the empirical situation occurs according to the theoretical distribution. It is given by:

\[
L_n(i) = \prod_{r=0}^{441} P_n(f(i, r)|s(r)) \tag{13}
\]
and the analogous equation for the cluster distribution (replacing ‘n’ by ‘c’). The maximum likelihood is the maximal value of $L_n(i)$ and $L_c(i)$, respectively, that can be reached for any parameter set. It is denoted by $\hat{L}_n(i)$ and $\hat{L}_c(i)$, respectively. The respective parameter sets determine those distributions that describe reality best. An example of a fitted natural and cluster distribution is given in Fig. 2 with the empirical distribution given in Fig. 1.

The aim of the study that is conducted here is to find out whether and in which industries local clusters appear. The additional term in the cluster distribution should describe such clusters. Thus, the cluster distribution should describe the empirical data significantly better than the natural distribution if clusters exist.

The cluster distribution contains six parameters, while the natural distribution contains only three parameters. Furthermore, the natural distribution is a special case of the cluster distribution. Therefore, the cluster distribution will always fit reality better, so that $\hat{L}_c(i) \geq \hat{L}_n(i)$ is satisfied for all industries.

The likelihood ratio test is used to check whether the cluster distribution describes reality significantly better than the natural distribution. This is done for each industry separately. To this end, the value:

$$\lambda(i) = 2\ln(\hat{L}_c(i)) - 2\ln(\hat{L}_n(i))$$  \hspace{1cm} (14)

is calculated. $\lambda(i)$ measures this difference in the fitting of the data. Statistical theories tell us that $\lambda(i)$ can be expected to follow a $\chi^2$-distribution if the additional term in the cluster distribution does not really make the cluster model more adequate than the natural distribution (Mittelhammer, 1995). Whether $\lambda(i)$ falls into this distribution can be tested. Hence, the hypothesis that the cluster distribution is not more adequate than the natural distribution can be tested. If it is rejected, the industry is said to be clustering.

The likelihood ratio test answers the question of which distribution describes the empirical data better. However, it does not answer the question of whether the distributions describe the empirical data adequately. To test the adequacy of the distributions, the distribution that is more adequate in each industry is compared with the empirical distribution.

To test whether the theoretical distribution and the empirical distribution deviate from each other significantly, the Kolmogorov–Smirnov test is used. The Kolmogorov–Smirnov test compares the cumulative distribution function of a theoretical and an empirical distribution. It makes a statement about the maximal distance of these two functions that should occur with a certain probability if the two distributions are identical. Hence, if the distance is too large, the hypothesis that the theoretical and the empirical distribution are identical can be rejected.

### Identification of local clusters

Finally, it is intended here to identify all local clusters in those industries that are found by the approach described so far to show clustering. Above it has been assumed and confirmed that regions containing a local cluster can be described by a Boltzmann–like distribution with a minimal number of $\xi_j$ (see equation 2). Hence, $\xi_j$ can be seen as the critical number of firms that constitute a cluster. It is formulated either in relative or absolute terms (see equation 8), depending on the definition of the size, $s$, that is used. In the case of relative numbers of firms, $s$ and therefore also $\xi_j$ differ between regions, while $s$ is the same for all regions in the case of absolute numbers of firms. This means that the critical value, $\xi_j$, is obtained as a result of fitting the theoretical model to the empirical data. Thus, information is obtained about how large clusters are.

This information is then used to identify the existing clusters in Germany. All regions in which the number of firms in an industry exceed the critical value $\xi_j$ are called local industrial clusters here. Five restrictions on the approach have to be mentioned in this context.

First, the approach used here is not able to distinguish between different causes of the existence of local clusters. If there are natural causes, such as the availability of necessary resources, for the existence of an agglomeration of firms, this agglomeration will be identified as a cluster here. However, this only happens if the distribution of firms among regions that is caused by the natural circumstances is in line with the predictions obtained through the assumption of local self-augmenting processes. Otherwise the empirical approach that is used here will not confirm the existence of local clusters in the industry under consideration.

Second, the two parts of the cluster distribution overlap to some extent in quite a number of cases. This implies that an exact distinction between regions containing a cluster and regions not containing a cluster is not feasible. However, in most cases the overlap concerns less than two regions. Nevertheless, it has to be kept in mind that the method used here might identify some additional clusters.
EMPIRICAL RESULTS

The method described above allows each industry to be examined to determine whether local clusters exist, how many firms have to be located in a region to constitute a cluster, and where the clusters are located. The method is applied to all three-digit industries in Germany. The results are given below. First, some general findings are reported. Then a list of all identified local clusters in the manufacturing sector in Germany is presented. Finally, these findings are compared with the existing case studies and are discussed from a regional perspective.

General results

According to the German WZ-93 classification there are 20 three-digit industries in the farming and mining sectors, 104 three-digit industries in the manufacturing sector, and 97 three-digit industries in the service sector. For each of these industries the methodology developed above is applied to the absolute and the relative numbers of firms. The test of the theoretical model shows that it can be rejected for none of the industries in the farming and mining sector. It can be rejected for seven industries in the manufacturing sector, but only with respect to the relative numbers of firms. The distribution of the absolute numbers of firms among regions is adequately described by the theoretical model in all manufacturing industries. In the case of the service sector, the theoretical distribution is rejected for both the absolute and the relative numbers of firms for 19 industries. In addition, for 29 industries the theoretical distribution is rejected for the relative numbers of firms.

Hence, the theoretical model fits the service sector quite badly. The reason for this is the fact that the theoretical model considers only stochastic influences and clustering forces. In addition, many service industries are characterized by forces that lead to a uniform distribution in space, caused by the fact that consumers and customers only approach local firms for the respective services. This holds especially for the more traditional service industries such as wholesale, retail, most business services, transportation, construction, and energy and water supply. These industries mainly serve customers from the region and do, therefore, compete for local markets. This causes a rather uniform spatial distribution of activities, except of some concentration in places where many people/customers are also located. Such forces are not represented by the model. However, there are also service industries such as software and media that have characteristics similar to manufacturing industries.

The further analyses are restricted mainly to the manufacturing sector because this sector is most discussed in the context of clustering, and the theoretical model seems to represent the situation in the manufacturing industries adequately. However, when identifying local clusters, the software and media industries are also included.

Table 1 shows that number of manufacturing industries for which local agglomerations are shown to exist. In total, local agglomerations are found for at least one of the two numbers of firms – relative and absolute – in 50 of the 104 manufacturing industries. Hence, clustering forces seem to be common but not universal in the manufacturing sector.

<table>
<thead>
<tr>
<th></th>
<th>Relative numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute numbers</td>
<td>Cluster</td>
</tr>
<tr>
<td>Cluster</td>
<td>17</td>
</tr>
<tr>
<td>Natural</td>
<td>17</td>
</tr>
</tbody>
</table>
The analysis conducted here shows that not all industries contain local clusters and that different critical values are used (ISAKSEN, 1996). Hence, for a few industries the value found herein fits the value assumed in the literature quite well. However, for many industries $\xi_i(i)$ is much higher than it is assumed to be in the literature. Hence, the number of local industrial clusters that exist is overestimated in the literature. Furthermore, such a critical value for the identification of local clusters is applied equally to all industries in the literature. The analysis conducted here shows that not all industries contain local clusters and that different critical values hold for all industries.

Identified local industrial clusters

For 54 manufacturing industries the existence of local clusters has been confirmed. In these industries a total of 383 local industrial clusters are identified according to the approach taken here. It is likely that these agglomerations are local industrial clusters. The values of $\xi_i$ range from 3 to 87. However, values between 3 and 20 are found most frequently. Values larger than 30 only appear in those cases where one region contains a very large agglomeration of firms and is identified as the only local cluster. The distribution of the value of $\xi_i$ found in the empirical analysis is shown in Fig. 3.

In the literature, for example, a value of $\xi_i(i) = 3$ is used (ISAKSEN, 1996). Hence, for a few industries the value found herein fits the value assumed in the literature quite well. However, for many industries $\xi_i(i)$ is much higher than it is assumed to be in the literature. Hence, the number of local industrial clusters that exist is overestimated in the literature. Furthermore, such a critical value for the identification of local clusters is applied equally to all industries in the literature. The analysis conducted here shows that not all industries contain local clusters and that different critical values hold for all industries.

In addition, Munich is regarded as an innovative milieu in the literature because it is strong in the vehicle, aerospace, electronic engineering, precision instruments/optics, and office machinery/data processing industries (STERNBERG and TAMÁSY, 1999).

The comparison leads to very different results. In some cases quite a good correspondence is found, although industries or regions are not always defined in exactly the same way. For example, the industrial districts in textiles show up in the analysis conducted here as a local cluster in textile processing, other textiles, clothes and curtains. Similarly, metal processing is subdivided into many different industries here. Many of them show clustering. Most of these local clusters are found in the Ruhr area. The same holds for the media industry. In some of these cases more local clusters are found in the above analysis than are described in the literature.

There are also many industries for which local clusters are discussed in the literature but no significant clustering is observed in the above analysis. These are the machinery-, automobile- and train-building industries. This discrepancy might have different causes. First, the
statistical approach used here is only able to reject the assumption that no local clusters exist in an industry. It might fail to identify all industries that show clustering. Second, the approach taken here is based on the distribution of firms, while case studies usually refer to the number of employees in an industry and region. For example, in the case of automobiles the clusters described in the literature are characterized by a very large factory site or company. The above approach does not identify the locations of large firms, but only agglomerations of many firms. Therefore, it might lead to very different results. Finally, in the literature regions are often discussed that are much larger than administrative districts. Examples are the machinery and printing machinery industry regions. If clustering occurs in such large spatial units, it is not identified by the approach taken here. However, it might be doubted whether, in such large regions, the local self-augmenting processes that are discussed in the theoretical literature have a significant impact on the developments.

The case of biotechnology differs from the other cases discussed. Biotechnology firms are usually assigned to the industries of pharmaceuticals or measurement instruments in the WZ-93 classification. Some even belong to other industries. Hence, the analysis cannot be given. This is done in Fig. 4. A similar picture is produced by Sternberg and Litzenberger (2004, figure 5). There can be similarities as well as differences observed between these to pictures. It is likely that these differences result to some extent from the different approach, but especially from the different aggregation level of the two analyses. A further analysis of the differences is not possible because Sternberg and Litzenberger do not provide a detailed list of the local clusters they identify.

The spatial distribution shown in Fig. 4 has several characteristics that are worth discussing here. First, 158 of the 441 administrative districts contain at least one industrial cluster. It has been stated above that the approach taken here might underestimate the number of industries that show clustering and, therefore, might underestimate the total number of clusters. Hence, a sizeable share of all districts contains industrial clusters. In other words, the existing local industrial clusters are well-scattered across space. Nevertheless, local industrial clusters are not everywhere. They are not uniformly distributed across Germany. There are some areas in which they are concentrated. In Fig. 4 three such areas can be easily identified: an area south of the Ruhr; an area in the north of Bavaria and the south of Thuringia and Saxony; and an area in the south of Baden-Württemberg. In addition, there are many smaller spots, many around large cities such as Berlin, Hamburg, Hanover, Leipzig, Frankfurt am Main, Stuttgart and Munich. Some areas are empty. This is especially the case in the north-east where, for example, no cluster is found in the state of Mecklenburg-Western Pomerania.

The large empty spots match quite well with those regions in which there is little industrial employment and development. These regions seem unable to exceed the critical mass for the emergence of local industrial clusters. Although random processes are involved in the determination of the location of clusters and although these clusters are quite scattered within Germany, certain regions do not seem to offer sufficient circumstances. This can be seen as evidence for the existence of at least one necessary condition with respect to local circumstances. However, it does not help one to identify this necessary condition. In this context one can only speculate. There is little discussion about what is lacking in those regions that are lagging behind (an exception is found in SERI, 2003). Research in this direction would be helpful to understand the local conditions necessary for the emergence of industrial clusters.

While the empty spots in Fig. 4 match the distribution of economic activity in Germany quite well, the same does not hold for the place in which many local clusters exist. In Baden-Württemberg (south-west Germany) there should be many more districts containing one or several clusters, for example in the east of Stuttgart and along the Rhine from Karlsruhe to Mannheim. At the same time the strong concentration of clusters in the north of Bavaria and the south of Thuringia does not match the lack of economic prosperity of this area. This shows that the existence of local industrial clusters should not be mixed up with economic prosperity. They might differ tremendously. Two factors are responsible for this. On the one hand, the spatial distribution of local clusters reflects the history of places rather than their actual situations. Many clusters remain stable over decades or even centuries. For example, the area in the north of Bavaria and the south of Thuringia and Saxony was strong in handicrafts many decades ago. Textiles, porcelain, toys, musical instruments and other similar things were produced there. This is still the case today. Although employment has decreased strongly in many of these industries, the respective districts are still strong in these industries compared with other regions in Germany. The same holds for metal processing and metal goods production south of the Ruhr. These clusters have formed quite some time ago. At that time they cause economic prosperity for the region. Later on they
remain stable in many cases, but without a further impulse for economic development, and sometimes even cause problems for the region if the respective industry declines (Grabher, 1993; Poudér and St. John, 1996).

In the history of Germany, mountain areas with a poor small-farm agriculture have been more likely to benefit from industrialization, so that local clusters in traditional industries are found in these regions. Local clusters in newly developed industries are rather found in big cities (see the list in the Appendix). On the other hand, the identification of local industrial clusters neither identifies regions that prosper because of a heterogeneous mix of industries nor identifies those characterized by a large, successful firm. Baden-Württemberg is usually claimed to be one of the major locations of the automobile, machinery and chemical industries (for a detailed discussion of the economic situation in Baden-Württemberg, see Cooke and Morgan, 1994). However, most of

Fig. 4. Clustering in Germany (shaded according to the number of industrial clusters that each district contains)
Identification of Local Industrial Clusters in Germany 1001

CONCLUSIONS

The above approach shows that the phenomenon of clustering can be studied empirically using a general approach. The main result is that clustering is more than a random result of an uneven distribution of local circumstances. There is evidence for the existence of local self-augmenting processes that cause the industrial firm distribution among regions to be multimodal. Based on this finding, those industries in which local clusters exist in Germany are identified herein. Furthermore, it has been shown that the condition with respect to the number of firms for the identification of local clusters can be obtained empirically. This means that the respective condition does not have to be set arbitrarily but it can be deduced from the empirical data for each industry separately. This makes the study of local industrial clusters more objective and helps to choose regions and industries for further case studies.

Acknowledgements — The author thanks Dirk Fornahl, Ulrich Witt, the participants of the ETE Workshop in Jena, Germany, an anonymous referee for helpful comments and discussions, and the German Federal Ministry for Education and Research for financial support. The usual disclaimer applies.

APPENDIX

For 54 manufacturing industries local clusters have been found to exist. A list of all identified local clusters is given below. The distilling industry is excluded because the number of firms is low so that all small regions that contain at least one firm show up in the list of local clusters. Some service industries are added to this list because they are repeatedly discussed in the literature. The list contains, for each industry studied, the names of all the administrative districts that satisfy the above condition (439 local clusters altogether).

For the following industries local clusters are found for the relative and the absolute numbers of firms (regions that contain clusters according to both numbers of firms are in bold):

- Fish processing: Dithmarschen, Nordfriesland, Schleswig-Flensburg, Cuxhaven, Wittmund, Bremerhaven (city), Wismar (city), Ostvorpommern, Rügen, Schöneberg.
- Textile weaving: Borken, Steinfurt, Hof.
- Clothing: Hamburg (city), Zollernalbkreis, Munich (city), Miltenberg, Berlin (West).
- Leather processing: Offenbach (city), Offenbach, Bad Kreuznach, Munich (city).
- Shoes: Pirmasens (city), Pirmasens.
- Sawmills: Hochsauerlandkreis, Ortenaukreis.
- Foundry: Solingen (city), Mettmann, Märkischer Kreis, Olpe, Pforzheim (city), Enzkreis, Aue-Schwarzenberg.
- Metal tools: Remscheid (city), Solingen (city), Wuppertal (city), Mettmann, Märkischer Kreis, Pirmasens (city), Eslingen, Rems-Murr-Kreis, Ostalbkreis, Berlin (West), Berlin (East), Schmalkalden-Meiningen.
- Lights: Hochsauerlandkreis.
Ships and boats: Ostholstein, Plön, Schleswig-Flensburg, Hamburg (city), Emden (city), Wilhelmshaven (city), Wesermarsch, Bremen (city), Bremerhaven (city), Berlin (West), Müritz.

Jewellery, coins and cutlery: Birkenfeld, Ostalbkreis, Pforzheim (city), Enzkreis, Kaufbeuren (city).


Film: Hamburg (city), Hannover, Düsseldorf, Cologne, Erfurtkreis, Frankfurt am Main (city), Wiesbaden (city), Stuttgart (city), Baden-Baden (city), Munich (city), Munich, Starnberg, Nürnberg, Berlin (West), Berlin (East), Potsdam (city), Dresden (city), Leipzig (city).

For the following industries local clusters are found for the relative numbers of firms:


Feeding grain: Verden, Cloppenburg, Vechta, Bördekreis.

Textile processing: Zollernalbkreis, Hof, Kulmbach.

Bed and decor textiles: Vogtlandkreis.

Other textiles: Wuppertal (city), Zollernalbkreis, Bayreuth, Hof, Neustadt a. d. Aisch–Bad Windsheim, Plauen (city), Annaberg, Chemnitz-Land, Vogtlandkreis, Mittweida.

Hosiery and knitwear: Mittlerer Erzgebirgskreis.


Basic chemicals: Ennepe-Ruhr-Kreis, Bitterfeld, Merseburg-Querfurt.


Weapons and munitions: Rottweil, Suhl (city).

Toys: Coburg, Mittlerer Erzgebirgskreis, Sonneberg.


For the following industries local clusters are found for the absolute numbers of firms:


Publishers: Hamburg (city), Düsseldorf, Cologne, Frankfurt am Main (city), Stuttgart (city), Munich (city), Berlin (West), Berlin (East).

Printers: Hamburg (city), Munich (city), Berlin (West).

Petroleum processing: Hamburg (city).

Tyres: Hamburg (city), Main-Kinzig-Kreis, Berlin (West), Berlin (East), Dresden (city).

Sculptor and stonecutter: Hamburg (city), Mayen-Koblenz, Berlin (West).

Ferrous metal production: Hagen (city), Märkischer Kreis, Siegen-Wittgenstein.

Non-ferrous metal processing: Märkischer Kreis.

Smiths: Märkischer Kreis.


Other metal goods: Märkischer Kreis.

Machine tools: Märkischer Kreis, Enzkreis.

Electrical equipment: Essen (city), Mettmann, Cologne, Dortmund (city), Märkischer Kreis, Esplingen, Karlsruhe, Berlin (West), Berlin (East).
• **Electronics**: Märkischer Kreis, Schwarzwald-Baar-Kreis, Dresden (city).
• **Telecommunication**: Düsseldorf, Frankfurt am Main (city), Munich (city), Berlin (West), Berlin (East).
• **Medical instruments**: Hamburg (city), Tuttlingen, Munich (city), Berlin (West).
• **Measurement instruments**: Hamburg (city), Lahn-Dill-Kreis, Munich (city), Berlin (West).
• **Aerospace**: Bodenseekreis, Munich.
• **Motorcycles**: Bielefeld (city).
• **Software**: Hamburg (city), Cologne, Frankfurt am Main (city), Stuttgart (city), Munich (city), Munich, Berlin (West), Berlin (East).
• **Data processing**: Hamburg (city), Hanover, Düsseldorf, Bonn (city), Cologne, Dortmund (city), Frankfurt am Main (city), Stuttgart (city), Munich (city), Berlin (West), Berlin (East), Dresden (city), Stuttgart (city), Munich (city), Leipzig (city).
• **Radio and television**: Kiel (city), Hamburg (city), Hanover, Düsseldorf, Bonn (city), Cologne, Frankfurt am Main (city), Mainz (city), Baden-Baden (city), Munich (city), Munich, Augsburg, Berlin (West), Berlin (East), Potsdam (city), Dresden (city), Leipzig (city).

**NOTES**

1. The exact form used is obtained in physics under certain conditions for the more general Boltzmann distribution. For simplicity it is called ‘Boltzmann distribution’ herein.
2. Industries are classified according to the WZ-93 classification, which is currently the standard classification of industries in Germany.

**REFERENCES**


