# Working Paper Series
## No. 17

# Identification of the Portuguese industrial districts

João C. Cerejeira da Silva

February 2002

**Núleo de Investigaçõo em Microeconomia Aplicada**
**Universidade do Minho**

Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA E DA TECNOLOGIA

# Identification of the Portuguese Industrial Districts

**João C. Cerejeira da Silva***

University of Minho - NIMA
Portugal

February, 2002



For any query or information, please contact the author:

jccsilva@eeg.uminho.pt

Abstract

Some authors have tried to define a methodology of identification of the local production systems, namely in terms of the operationalization of the notion of "industrial district". For the Portuguese case, there is no previous work, using of a systematic methodology of the identification, on the identification of the industrial districts, in spite of the existence of some case studies.
In this paper we propose an algorithm of classification, based on the cluster analysis, and we try to find clusters of homogeneous geographical units, in order to identify the ones that we might classify as industrial districts.
Our results point that almost one third of the Portuguese employment in manufacturing and 13% of all employment, is located in industrial districts. A detailed analysis of other variables, shows that the Portuguese industrial districts' characteristics are very close to the ones found in other contexts.

* I would like to thank Octávio Figueiredo and Paulo Guimarães for helpful comments and help. Financial support by the Portuguese Foundation for Science and Technology (FCT) and University of Minho is gratefully acknowledged. The usual disclaimers apply.

## 1. Introduction

Several authors have tried to define a methodology to identify the local production systems, regarding the concept of “Industrial District”. Among the most known articles in literature, two important contributions have to be mentioned - those ones from the pioneers Garofoli (1983 and 1994) and Fabio Sforzi (1987 and 1990). On the first case, the author searches for areas of manufacturing specialization characterized by small companies within a small geographic area, in order to classify those areas into three different groups, in accordance with their systemic nature: areas of manufacturing specialization, local production systems and system-areas.

On the second case, the main goal is to identify the Marshallian industrial districts in Italy, starting from the collection of the 955 Local Labour Market Areas (LLMAs). These areas are space-functional units, defined through the functional geography of the house/work flow. Sforzi (1987) uses the multivariate analysis and classifies the LLMA into 15 categories, one of them including 61 industrial districts, almost all located in the Northeast and Centre regions of Italy.

Regarding Portugal, besides all the studies made on the subject, there is no knowledge of other works of systematic classification in a “concelho” scale, a small administrative area that would categorize the local productive or industrial systems in a national scale. We could mention however the analysis of the coastal northern region configuration (48 “concelho”, listed in Silva (1988) and Silva e Figueiredo (1992), where we can find a group of local productive areas with different industrialization and complexity levels.

Our work intends to carry out the segmentation of the entire country into “concelho” groups, slightly homogeneous, regarding certain pre-defined characteristics, in order to identify those that might be regarded as industrial districts. We will use as methodology, multivariate statistics proceedings and, as classification technique we will use the cluster analysis with an appropriate variant for the spatial analysis. The application of multivariate statistical methods for the geographical classification is not very common in Portugal, in spite of the existing works for other countries. In fact, we can only find the works of Brandão, Pires and Portugal (1998), as well as the works of Gomes, Bacelar and Saleiro (1994) for the Northern Region. All those works make use of the information provided by the Population Census of 1991, but strictly for the purpose of description of the resulting regions, concerning the social, economical and political aspects.

Our results point that almost one third of the Portuguese employment in manufacturing and 13% of all employment, is located in industrial districts. A detailed analysis of other variables, shows that the Portuguese industrial districts' characteristics are very close to the ones found in other contexts

The structure of the paper is as follows: Section 2 presents the clusters analysis as a method of spatial classification, Section 3 describes the application of the method and shows the main results, and, finally, Section 4 concludes.

## 2. The cluster analysis used as a classification method for spatial analysis

The importance for classification seems very obvious in the knowledge areas where the categorization of the elements is a structural basis; such is the case of biology (i.e. taxonomy of the species) or medicine (i.e. classification of diseases). Generically speaking, the cluster analysis is a set of multivariate statistic methods that give the possibility to assemble a group of $N$ individuals, categorized by $q$ variables, in $K<N$ groups relatively homogeneous, generically named as clusters, regarding only the similarities and differences among them[1]. In social sciences, we can find some efforts towards the application of the cluster analysis in so varied areas such as the case of anthropology, political science or even business-related sciences[2]

Concerning the spatial analysis, there is enumerable situations where classification or taxonomy of the basic spatial units into categories, according to pre-defined characteristics, can be useful. Beguin (1979) lists five goals where the classification of spatial units may apply:

(1) The goal of simply reducing the number of units for analysis;

(2) The goal of defining homogeneous regions, i.e. groups of continuous regions with similar features;

(3) The goal of defining a type system for the spatial units, even if they are not continuous regions (as in the previous case);

(4) The goal of generate exploratory hypothesis for future research;

(5) Finally, the goal of testing the theories;

[1] The most valuable contribution for the application of these methods was given by Robert Sokal and Peter Sneath in 1963, in the book: *Principles of Numerical Taxonomy*, oriented specifically for the biological classification

Naturally, the different categorization criteria also mean different results. Cliff et al. (1975) suggest that an optimum categorization of geographical units should respect the following criteria:

a) They must be straightforward, in order to provide a solution which produces the minimum number of categories as possible;

b) The spatial units classified into the same category must be more or less similar, in respect of their variables;

c) If the goal is the creation of new regions, the principle of adjacency among units under the same category has to be preserved.

**2.1. The cluster analysis**

In short, the cluster analysis tries to group the several elements by using the existing information and always focusing on the similarities between the constituents of the same group, which have to be always more noticeable then the similarities between different groups. In spite of the several criteria for dividing and grouping the objects, the different goals for the analyses or even the different types of initial data, there are five main stages in clusters that have to be respected:

(1) The selection of the sample of the objects to be grouped.

(2) The definition of set of variables that characterize the sample objects.

(3) The selection of a similarity or dissimilarity measure for each pair of objects.

(4) The selection of a classification algorithm and respective application.

(5) Finally, the validation of results.

We will discuss further on the main aspects of each different stage listed above.

**2.2. The selection of the variables that characterize the sample of units to be grouped**

The selection of the objects to be grouped and their own variables lead us to a double problem: On one hand, it is necessary to chose out the available objects, those to be used in the classification method; but this problem will be solve depending on the information the investigator has on the objects, at earlier stage. On the other hand, there is a statistical problem regarding the possible application of variables that are defined with different scales.

[2] As examples of application of clusters analysis in the entrepreneurial sciences, we have the market segmentation or the identification of company groups with similar strategic behaviours (see Reis 1997).

In spite of a certain controversy, regarding the procedures for such case[3], if the distribution is normal and the variables are independent between themselves, the procedure most commonly used is that one of the standardization of the variables, through their conversion into new variables of type $Z = (X - \mu)/\sigma$, where $X$ represents the original variable; $\mu$ represents the sample average of the variable and $\sigma$ represents its standard deviation.

However, we have to bear in mind that if this standardization process imposes a null average and a unit standard deviation for all converted variables, it will consequently reduce the differences between the individuals as well as the importance of the variables.

In the case where there area several variables bearing a major relevance, is possible to give them more weight, especially if there are strong theoretical reasons behind it.

**2.3 Selection of similarity and dissimilarity measure for each pair of objects**

The (dis)similarity measures normally used in social sciences are the same used to measure the difference or distance between the several elements of a data matrix. Although several metrics are possible, the measure most commonly used between two cases $i$ and $j$ is the *Euclidian Distance* ($d_{ij}$): defined as the square root of the summation of the differences between the values $i$ and $j$ for all variables ($v = 1, 2, \ldots, q$),

$$d_{ij} = \sqrt{\sum_{v=1}^{q} \left(x_{vi} - x_{vj}\right)^2}.$$

The application of this metric requires the standardization of the variables, which can raise some problems. Besides, the results obtained by any classification algorithm will be forcibly dependent on the employed metrics.

**2.4. Identification of clusters on the spatial analysis**

Although there is no exact definition for the word cluster, the proposal of Everitt (1980) seems to be generally accepted: “Clusters are continuous regions of [a] space containing a relatively high density of points”. Consequently one of the problems arising in any analysis is the need to select a criterion of (dis)aggregation of the several cases with the purpose of creating groups whose elements are similar between themselves.

There are several “families” of clustering techniques, each one representing a different perspective of group formation, which is reflected on the different classification algorithms

[3] See Aldenderfer and Blashfield (1984, p. 20)

that are used. This way, it may happen that the obtained results are not identical, when different methods are used for the same object sample. Although there is no such thing as the best clustering method, and since all methods have advantages as well as disadvantages, some particular techniques are used in a more regular way in order to solve specific problems.

In the case of spatial analysis, the most recent works indicate the application of optimization methods (Wise, Haining and Ma, 1997; Murray, 1998), with detriment to hierarchical techniques, being the *k-means* method (MacQueen, 1967) the best-known example.

There are currently three optimization models that may be used in the resolution of spatial classification problems[4]. The best-known and more often used method is the CPCP[5], which is a variant for the spatial analysis of the *k-means* method (MacQueen, 1967), based on the works of Cooper (1963) that can be found in the majority of the statistical packages. This method consists in grouping the spatial observations by minimizing their Euclidian distance in relation to central points that are artificially created. We may, therefore, use the following notation:

$i$ = localization index of each one of the $n$ observations ($i$ = 1, ..., $n$), characterized by a vector **V** of variables;

$p_p$ = central point (centroid) of cluster $p$ ($p$= 1,…$k$);

$d_{ip}$ = Euclidian distance between observation located in $i$ and $p_p$.

$$y_{ip} = \begin{cases} 1 \text{ if observation located in } i \text{ belongs to cluster } p \\ 0 \text{ otherwise} \end{cases}.$$

The purpose of CPCP is minimizing the total difference between the coordinates of each observation (variable observed values) and the coordinates of the central points:

$$\min Z = \sum_{i=1}^{n} \sum_{p=1}^{k} d_{ip} y_{ip}$$

with the condition that all observations are connected to one cluster only:

$$\sum_{p=1}^{k} y_{ip} = 1.$$

In the case of CPCP method, the best known and more often used heuristic, is as follows (Murray, 1998):

1st step: Selection of an initial partition of the spatial units by a $k$ number of clusters, designated by the researcher.

---

[4] For a recent revision of these models, see Murray and Estivil-Castro (1998).

2nd step: Calculation of respective centroids

3rd step: Calculation of distances between each observation and the centroids of the different groups, transferring each individual to the group located at a lower distance.

4th step: If, in the previous step, there was not any transfer of observations from one cluster to another, the heuristic ends, which means a local optimum solution was found. Otherwise, it is necessary to go back to the 2nd step.

One of the major drawbacks of these techniques is the fact their performance depends on the selection of an initial partition, in order to avoid a sub-optimum solution[6]. However, Milligan (1998) has demonstrated that the *k-means* method, when using an initial partition resulting out of the application of a hierarchical technique, obtains a higher performance than the hierarchical methods, in terms of obtaining a global and not a local optimum solution.

### 2.5. Validation of results

Since the purpose of the clusters analysis is creating homogeneous groups, the decision on the number of clusters become a problem itself. In fact, there is no previous knowledge about the number of groups in which the object population is being divided into. One of the methods to be used in hierarchical techniques is the chart comparison of the number of clusters with the respective fusion coefficient, that is, the numerical value (similarity or dissimilarity) to which different objects are aggregated to create a group. In that sense, when the division of a certain group does not cause significant changes on the fusion coefficient, it may become optimum.

Another existing procedure is also the comparison of the results, by using other different grouping criteria. It is possible to obtain the convergence degree between the several grouping criteria, by means of a contingency table, indicating the number of observations, which are grouped under the same cluster, for the same number of clusters. That way, it is possible to check the level of stability of the existing solutions, in order to come to a conclusion about the quality of the grouping.

---

[5] CPCP – Centre Points Clustering Problem.

[6] The possibility of the solution being a local and not global optimum solution, results directly from the functional incapacity for generating all possible partition combinations of the *n* observations in *k* groups, in order to choose the one that minimizes the sum of the distances between observations and

## 3. The clusters analysis used for identification of the industrial districts in Portugal

For the identification of the industrial districts in Portugal we will have four different main stages. First, it is necessary to choose an operational concept for industrial districts, i.e. a concept, which might be characterized by quantitative variables. The second stage consists of selecting the variables, on a certain analysis scale, as well as their quantification. The third stage covers the application of a classification technique and further selection of the group(s) whose characteristics should be the closest as possible to the industrial district concept that was previously suggested. The fourth and fifth stages are reserved for the validation of results, obtained through the identification of the industrial districts covered by the groups, with special regard to their specialization profile and the use of other complementary characterization variables.

### 3.1. The industrial district concept

Regarding the operational definition for industrial district, it is important to bear in mind the concept developed by Bianchi (1998), where he proposed a definition based on the simultaneous performance of three independent models: the *production model*, ruled by the flexibility characteristic, in a large sense, which is entered by a productive organization dominated by small companies, specialized in a specific product component or phase, operating in the same manufacturing line. Related to that specialization, we find a spatial concentration of several economical agents, (*spatial agglomeration model*), which gives the possibility to reduce transaction costs, training and labor qualification costs. This causes the creation of a solid and complex net of relationships between agents, supported by a strong cultural homogeneity and social consensus, which enables the local regulation of the economic relationships (*social model*).

### 3.2. The variables

Secondly, in order to identify the industrial districts in Portugal, it is necessary to create a typology for the spatial units (in our case, the 275 “concelhos” of the mainland), using a set of variables that may catch the most fundamental characteristics of the industrial district model. In our case, there was the selection of four variables that show four of those characteristics, namely, the industrial profile and the specialization of the “concelhos”, the concentration of

centroids of each group. Therefore, most of the optimising methods are heuristic and there are several heuristic alternatives.

workers in small and medium firms, spatially agglomerated, and finally, the geographical agglomeration of establishments of the same industrial branch. We must bear in mind that these variables basically characterize the production and spatial agglomeration model, which we refer to in the last paragraph. The variables, representative of the social model, will be discussed on the last stage to support the validation of results.

The selected variables, summarized in the next table, have “concelho” as their analysis unit (small geographical unit for which it was possible to obtain the necessary statistic data), and are as follows:

- The industrialization rate (PEMPIT), which is calculated as the employment share in the manufacturing industry, regarding the number of employed population in the local area;
- The manufacturing specialization coefficient (COEFESP); an indicator that compares the industrial structure of each “concelho” to the structure of a group used as reference (in this particular case, the mainland);
- The density of employment in small and medium firms (PME50KM), calculated by the number of workers in industrial units with less than 50 employees, per Km2;
- The industrial agglomeration index (ESPINDKM), calculated by the number of industrial units of the most important business branch of the “concelho”, per Km2;

Since all the variables belong to different scales, there we standardized them beforehand, through the transformation already mentioned.

*Table 1: Variables Description*

| Variable | Description | | Date | Fonte |
|---|---|---|---|---|
| PEMPIT | Industrialization Rate | $\frac{X_i}{PE_i}$ | $\frac{\text{Average1990\_92}}{1991}$ | $\frac{\text{Q.P.}}{\text{I.N.E.}}$ |
| COEFESP | Manufacturing Specialization Coefficient | $\sum_j \left( \frac{X_{ij}}{X_i} - \frac{X_j}{X} \right)^2$ | Average1990_92 | Q.P. |
| PME50KM | Density of Employment in Small and Medium Firms | $\frac{\overline{X}_i}{A_i}$ | Average1990_92 | $\frac{\text{Q.P.}}{\text{I.N.E.}}$ |
| ESPINDKM | Industrial Agglomeration Index | $\frac{Emax_i}{A_i}$ | Average1990_92 | $\frac{\text{Q.P.}}{\text{I.N.E.}}$ |

Notes and sources:

$X_{ij}$ = Employment in the industrial sector *j*, in "concelho" *i*.

$X_i$ = Total employment in manufacturing, in "concelho" *i*.

$PE_i$ = Employment in "concelho" *i*.

$X_j$ = Employment in the industrial branch *j*, in Portugal (mainland).

$\overline{X}_i$ = Employment in firms with less that 50 workers, in "concelho" *i*.

$Emax_i$ = Number of industrial units of the most important branch, in "concelho" *i*.

$A_i$ = Area of the "concelho" i.

Q.P. = "Quadros de Pessoal" of the Ministry of Labor and Solidarity

I.N.E. = Instituto Nacional de Estatística.

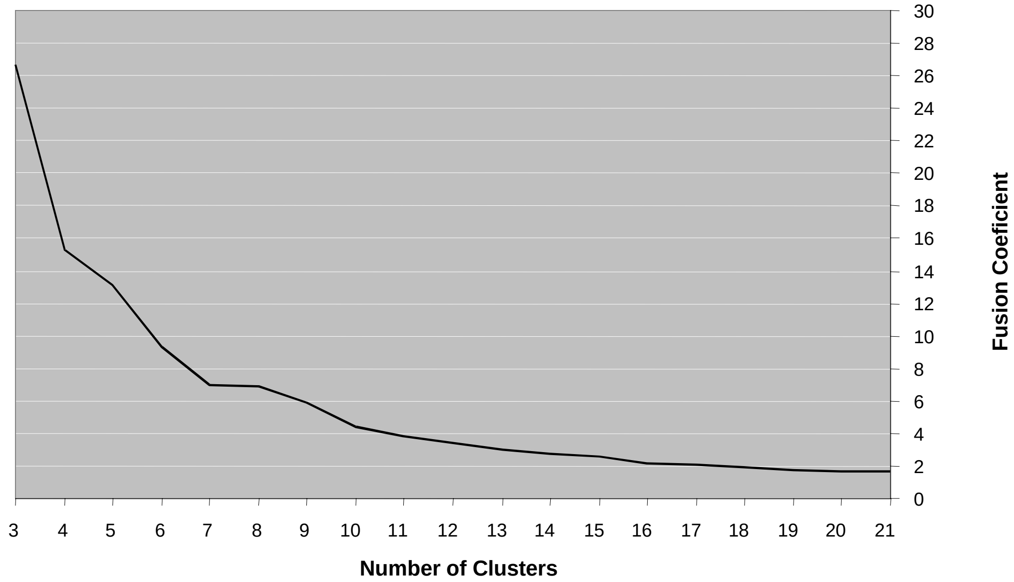
### 3.3. Classification using the cluster analysis and the results

As previously suggested, the CPCP was the classification method that was used, in its combining and inclusive variant. That is, the cluster centroid is recalculated every time there is a changing in its composition, including (or excluding) the coordinates of the observation that was included in (or excluded from) the cluster to which refers to.

The decision on the number of groups was made after a previous classification that was done with a hierarchical method, according to the inter-groups average criterion (see Reis, 1997, p. 317). By the simple observation of Graphic 1[7] that relates the number of cluster to the fusion coefficient, it is possible to accept the hypothesis of division of the "concelhos" into seven groups (fusion coefficient equal to 6,94), and then extract the respective centroids, in order to use them as initial centers on the implementation of optimizing method.

[7] In order to provide a better understanding of the graphic, we display only the most relevant part

**Graphic 1: Graphic analysis of the Fusion Coefficient**




On table 2, we show the final cluster centers, as well as the number of elements of each group. On the next chart, it is possible to observe the distribution of the concelhos by the several clusters and to obtain a global and more elucidative vision of the results.

*Table 2: Centers of the Final Clusters*

| **Variable (standardized)** | ***Cluster*** | | | | | | | **Median** |
|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | |
| Z(PEMPIT) | 2,30 | -0,51 | 7,08 | 0,49 | 0,19 | 2,06 | 1,29 | -0,36 |
| Z(COEFESP) | 0,15 | 0,51 | -0,13 | -1,25 | -0,79 | 0,01 | 0,82 | -0,25 |
| Z(PME50KM) | 0,27 | -0,22 | 11,80 | 6,70 | 0,00 | 7,86 | 0,82 | -0,22 |
| Z(ESPINDKM) | 0,31 | -0,19 | 13,94 | 3,33 | -0,06 | 5,35 | 2,72 | -0,20 |
| N. of elements | 20 | 151 | 1 | 2 | 99 | 1 | 1 | |

When we use the median of each variable as a comparing measure, because it is a location measure stronger than average, we may generically characterize the groups in the following way:

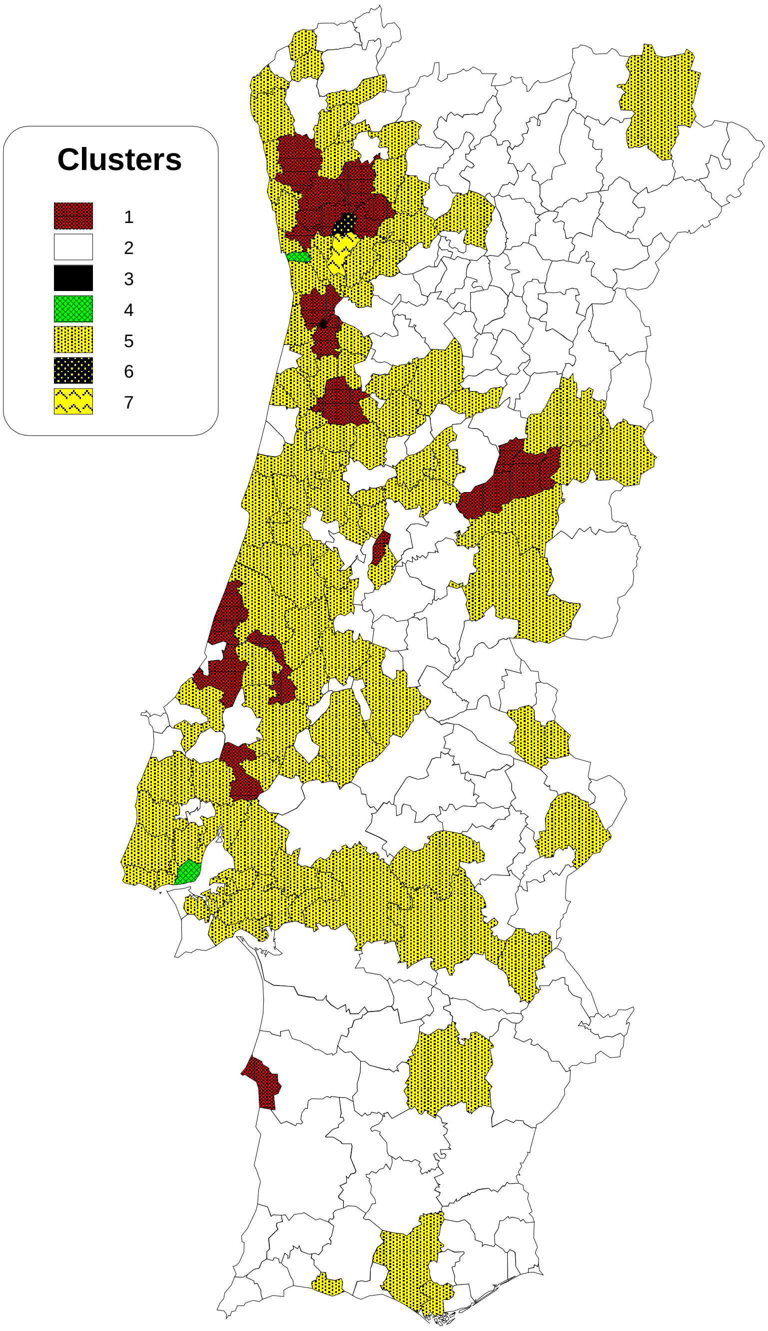
(1): Clusters 1,3,6 and 7 – In this set, the 4 clusters have centroids with values, which are higher than median, in all studied variables. The 23 concelhos, that compose these clusters, are located in four different areas: in the northern part of a ring, around the Oporto metro area ("Vale do Ave", part of "Vale do Cávado" and "Vale do Sousa"); in the Southern part of the same ring (Region of "Entre-Douro-e-Vouga"); in the involving the area of "Serra da Estrela", and finally, the region of "Norte da Estremadura".

In general terms, we may state that the clusters are characterized by a high industrialization rate, linked to a population of spatial agglomerated small and medium firms, and also by a deeply industrialized industry structure and a great number of industry establishments of the same branch, per Km2. We can, therefore, confirm that this cluster group holds the characteristics, which are closest to the industrial district model. It is indeed, within this group of "concelhos", that we will be able to identify the Portuguese industrial districts.

(2) Cluster 2 – Is the group with a higher number of elements (151). It comprises "concelhos" with a low industrialization rate and a low concentration of industrial units. As these "concelhos" have an insignificant number of industrial establishments, it is not surprising that the Specialization Coefficient may reach high values, since this indicator refers to the intra-industry specialization. In geographical terms, the "concelhos" of this cluster are mainly located in the Inland and in the South of the country, i.e. in regions where the primary and tertiary sectors are dominant.

(3) Cluster 4 – This cluster covers the "concelhos" of Lisbon and Oporto and they are characterized by a high concentration of industrial establishments and by a significant number of small and medium companies. The industrial structure is diversified and its industrialization rate is lower than the average in clusters 1, 3, 6 and 7.

**Map 1: Distribution of the mainland “concelhos” by the several clusters**



### 3.4. Validation of results

After the selection of the clusters, whose characteristics are better adjusted to the proposed definition of industrial district; it is necessary to select which of the "concelhos" may be identified as "centers" of the industrial districts. We performed that selection by removing from the group of 23 concelhos belonging to clusters 1, 3 6, and 7, the "concelhos" that show a value lower than a standard deviation, compared to national average, which means, it does not agree with all characteristics enumerated by the industrial district concept (only the "concelho" of "Maia" with the value of -1,31 in the variable COEFESP, is not covered by this criterion, but its industrial diversity keeps it out of the classification as industrial district). In the same way, the "concelhos" with less then 20 industrial units of the most representative sector were also excluded. That shows the lack of "critical mass" regarding the number of existing companies ("concelhos" "Belmonte", "Manteigas", "Batalha", "Castanheira de Pêra", "Azambuja" and "Sines")[8]. The table 3 shows the remaining 16 "concelhos" with respective employment rates (year of 1991) as well as the employment and the number of manufacturing establishments (average of years 1990, 1991 and 1992).

A global analysis of this table demonstrates that the employment of the 16 studied "concelhos" is about 13% of the total employment in Portugal (mainland), less then its share in the employment at manufacturing which represents almost one third of employment in Portugal (mainland) in the same business area. It is important to stress that all these "concelhos" have an employment rate of in the manufacturing industry that is higher than the average. We need to keep in mind that the average ratio for these "concelhos" (57%) is twice than the average in the mainland (24%). We are facing indubitably a high industrialization rate where the employment proportion, in that region and in that business area, is sometimes more than 50% of the total employment.

In addition, we detect a significant number of industrial establishments, never less than 200 and sometimes above 1000 in some "concelhos", which proves the existence of a great number of industrial units in each "concelho". On the other hand, all the "concelhos" show, in terms of geographic agglomeration, a spatial density of workers in the industry, above the average of the reference aggregate (10,7 workers per Km2, regarding the average of 86,5 per Km2 of the 16 concelhos).

[8] The lack of a significant number of manufacturing units in the most representative sector, does not mean the *concelho* belongs to a expansion or transition area, or to an industrial district

*Table 3: Selected "concelhos"*

| Concelhos | Total Employment (a) | Emloyment in Manufacturing (b) | Employment Share in Manufacturing (a)/(b) | Number of Manufacturing Establishments (c) | N. of workers in Manufacturing by KM$^2$ (d) |
|---|---|---|---|---|---|
| Águeda | 23.340 – 0,54% | 14.483  1,53% | 67,9% | 417 | 43,10 |
| Feira | 57.118 – 1,45% | 29.024 – 3,07% | 50,8% | 1.261 | 137,55 |
| Oliv. de Azeméis | 33.260 – 0,84% | 17.598 – 1,86% | 52,9% | 847 | 115,02 |
| Barcelos | 51.467 – 1,30% | 22.143 – 2,34% | 43,0% | 1.115 | 60,50 |
| Guimarães | 79.164 – 2,01% | 51.972 – 5,50% | 65,7% | 1.140 | 201,44 |
| V.N. Famalicão | 57.367 – 1,45% | 36.450 – 3,85% | 63,5% | 981 | 174,40 |
| Covilhã | 21.142 – 0,54% | 8.609 – 0,91% | 40,7% | 214 | 15,68 |
| Alcobaça | 24.091 – 0,61% | 10.770 – 1,14% | 44,7% | 365 | 26,08 |
| Marinha Grande | 14.049 – 0,36% | 9.944 – 1,05% | 70,8% | 241 | 53,46 |
| Felgueiras | 24.512 – 0,62% | 17.019 – 1,80% | 68,4% | 462 | 146,72 |
| Lousada | 20.321 – 0,52% | 7.980 – 0,84% | 39,3% | 282 | 79,01 |
| Santo Tirso | 51.899 – 1,32% | 33.250 – 3,52% | 64,1% | 700 | 159,86 |
| Alcanena | 6.305 – 0,16% | 4.510 – 0,48% | 71,5% | 203 | 35,51 |
| S. J. da Madeira | 9.441 – 0,24% | 12.806 – 1,35% | 35,6% | 336 | 1.829,43 |
| Paços de Ferreira | 21.823 – 0.55% | 11.011 – 1,16% | 50,5% | 886 | 161,93 |
| Paredes | 31.945 – 0,81% | 11.950 – 1,26% | 37,4% | 1.029 | 76,60 |
| Total Districts | 525.234 – 13,31% | 299.519 – 31,67% | 57,0% | 10.478 | 86,47 |
| Total Portugal (mainland) | 3.945.520 – 100% | 945.745 – 100% | 24,0% | 36.956 | 10,66 |
| **Sources:** | (a) - INE - CENSOS 91<br>(b) e (c)  - MTS - Quadros de Pessoal - averages 1990, 1991 e 1992. | | | | |

In short, table 3 shows very clearly that the all the presented "concelhos" are characterized by a significant number of industrial plants, spatially agglomerated. Regarding the intra-industry specialization of each "concelho", table 4 shows the first and second most dominant industrial sectors, in terms of employment volume and weight.

We have remarked that, the two most important industrial branches of each "concelho" cover more than 70% of the respective employment in the Manufacturing Industry (M.I.). It is also worthy of note, that the weight of both sectors is never below 50% of the employment in M.I. of each "concelho", with Covilhã showing the higher value (91%). In six different cases, we come across with clear situations of mono-specialization, i.e. complementary branches belonging to the same productive line (textile-clothing, or metal-transport material goods). In other cases, the specialization occurs in industry branches with similar features regarding the product type and the manufacturing process organization (the binomial footwear-clothing, for ex.). It is important to note, however, that the complementarities between dominant sectors (footwear / wood and cork, in the "concelho" de St$^a$.Maria da Feira, or wooden furniture - clothing, in Paredes and Paços de Ferreira, probably due to the closeness to Vale do Ave, or even ceramics – footware in Alcobaça), will be apparently reduced in 7 concelhos. Therefore, we came to the conclusion that, in spite of the strongly specialized industry structure, that specialization is not always based on the supremacy of a certain industry branch, since there is a dual specialization in approximately half of the studied “concelhos”. It means that we are dealing with industrial structures of greater complexity.

Table 4: Dominant Sectors and respective weight in the Manufacturing Employment

| Concelhos | Dominante Sector (a) | 2nd Dominant Sector (b) | Employment in (a) and % Regarding Total Emp. in Manufacturing | Employment in (b) and % Regarding Total Emp. in Manufacturing |
|---|---|---|---|---|
| Águeda | 381 – Metallic Products | 384 – Transport Goods | 5.108 – 35,3% | 2.059 – 14,2% |
| Feira | 324 – Footwear | 331 – Wood and Cork | 10.165 – 35,0% | 9.829 – 33.9% |
| Oliv. de Azeméis | 324 – Footwear | 322 – Clothing | 8.476 – 48,2% | 2.196 – 12,5% |
| Barcelos | 321 – Textile | 322 – Clothing | 11.020 – 49,8% | 6.888 – 31,1% |
| Guimarães | 321 – Textile | 322 – Clothing | 29.269 – 56,3% | 9.974 – 19,2% |
| V.N. Famalicão | 321 – Textile | 322 – Clothing | 19.416 – 53,3% | 8.113 – 22,3% |
| Covilhã | 321 – Textile | 322 – Clothing | 5.417 – 62,9% | 2.436 – 28,3% |
| Alcobaça | 361 – Ceramics | 324 – Footwear | 4.025 – 37,4% | 1.397 – 13,0% |
| Marinha Grande | 362 – Glass | 381 – Metallic Products | 4.006 – 40,3% | 1.697 – 17,1% |
| Felgueiras | 324 – Footwear | 321 – Textile | 12.886 – 75,7% | 1.305 – 7,7% |
| Lousada | 322 – Clothing | 324 – Footwear | 5.099 – 63,9% | 1.029 - 12,9% |
| Santo Tirso | 321 – Textile | 322 – Clothing | 15.942 – 47,9% | 9.981 – 30,0% |
| Alcanena | 323 – Leather | 321 – Textile | 2.607 – 57,8% | 1.092 – 24,2% |
| S. J. da Madeira | 324 – Footwear | 382 – Non Elect. Machines | 6.636 – 51,8% | 1.048 – 8,2% |
| Paços de Ferreira | 332 – Wooden Furniture | 322 – Clothing | 5.335 – 48,5% | 3.758 – 34,1% |
| Paredes | 332 – Wooden Furniture | 322 – Clothing | 7.839 – 65,6% | 2.152 – 18,0% |
| Total Distriticts | | | 153.246 – 51,2% | 64.954 – 21,7% |
| **Sources:** | MTS - Quadros de Pessoal - média dos anos de 1990, 1991 e 1992. | | | |

The dominant sectors of these “concelhos”, show a almost perfect adjustment concerning the specialization profile of the Italian industrial districts, identified by Sforzi (1990): “The majority of [Italian] Marshallian industrial districts have a dominant manufacturing specialization in fashion wear industries – i.e. textile, clothing footwear, leather goods and tanneries - and wooden furniture; a lower number is dominated by metal goods industries, mechanical and electrical engineering.” p.84.).

The conclusion that the above mentioned “concelhos” are industrial districts, would be not completed without making an additional reference to local social system, because for each manufacturing system there is a corresponding territorial back-up context which builds it up and influences its evolution. In that sense, both, social and labor force reproduction, have to be taken into account beyond the economy context, with special regard to the meaning of the cultural and symbolic interferences on the modeling of material relationships (Reis, 1992. p.229).

In the specific case of an industrial district formation, there are several determinants as the social homogeneity flexibility, that ensure, in one hand, the system reproduction (of labor, of entrepreneurial skills, …) and on the other hand, its own cohesion and local regulation of class conflicts. In the Portuguese case, the issues related to the territorial occupation and organization (demographic related) and the nature of small agriculture, support those two determinants. On the first one, several authors defend as condition for bringing up an industrial district, the existence of a significant density of human resources in order to ensure the labor force availability and a considerable communication network, associated to an urban system based on small and medium cities which offers several tertiary functions, avoiding at the same time, the agglomeration des-economies of the big cities (not only regarding the environment degradation, but specially the cost of certain factors as land or labor).

The other determinant is the agricultural structure based on the role of the owner's family and in small units (Reis, 1985, 1992; Silva, 1988). Several elements that characterize the rural and family-based society, become important for structuring and strengthening the industrialization model, specially regarding its task to render the employment market more flexible (with the contribution of home work and multi-activity forms), for maintaining a work ethics (expressed in considerable activity rates), and, for the entrepreneur factor it-self, resulting from the intimacy with the management and economical calculation functions.

In spite of the reduced available information, the indicators shown in table 5 intend to exemplify in a resumed way, the aspects above mentioned.

*Table 5: Several elements that characterize the Social Model*

| **Concelhos** | **Activity Rate (a)** | **Female Activity Rate (b)** | **Average Dimension of the Farms (c)** | **Av. Number of Elements in Each Household (d)** | **Population Density (d)** | **Urbanization Rate (e)** |
|---|---|---|---|---|---|---|
| Águeda | 49,4% | 41,4% | 0,75 | 3,30 | 131 | 22,3% |
| Feira | 49,5% | 40,9% | 1,51 | 3,50 | 562 | 39,5% |
| Oliv. de Azeméis | 50,8% | 41,8% | 1,30 | 3,50 | 437 | 38,8% |
| Barcelos | 47,3% | 44,0% | 2,38 | 4,00 | 305 | 7,8% |
| Guimarães | 52,1% | 44,2% | 2,75 | 3,80 | 611 | 21,3% |
| V.N. Famalicão | 51,8% | 43,5% | 3,10 | 3,50 | 547 | 23,9% |
| Covilhã | 42,4% | 39,2% | 7,23 | 2,80 | 98 | 23,6% |
| Alcobaça | 45,7% | 37,9% | 2,24 | 3,00 | 132 | 32,7% |
| Marinha Grande | 45,9% | 37,6% | 1,43 | 3,00 | 173 | 100,0% |
| Felgueiras | 49,1% | 41,2% | 2,14 | 3,70 | 442 | 13,3% |
| Lousada | 48,9% | 40,3% | 2,47 | 3,80 | 421 | 0,0% |
| Santo Tirso | 52,9% | 42,7% | 2,94 | 3,40 | 493 | 38,7% |
| Alcanena | 45,2% | 37,9% | 3,48 | 2,90 | 113 | 0,0% |
| S. J. da Madeira | 52,8% | 43,7% | 1,73 | 3,50 | 2.636 | 100,0% |
| Paços de Ferreira | 50,6% | 38,3% | 1,78 | 3,90 | 650 | 13,7% |
| Paredes | 45,6% | 32,9% | 1,97 | 3,80 | 468 | 40,2% |
| Average Districts | 49,5% | 41,4% | 2,24 | 3,58 | 318 | 28,5% |
| Average Portugal | 44,9% | 37,8% | 7,04 | 3,08 | 106 | 53,8% |
| **Notes:** | **Urbanization Rate**: % of population living in villages with more than 5.000 inhabitants. | | | | | |
| **Sources:** | (a), (b), (d), (e), e (f): INE – CENSOS 91.<br>(c) – INE – 1989 | | | | | |

Regarding the territory occupation, we emphasize the fact that most of the "concelhos" associate a high population density, above national average (the only exception is Covilhã, although close to the average), to a low[9] urbanizing rate. These indicators show that territories, which are not entirely rural or urban, have a high industrialization level as a result of the people dispersion (in spite of the intense communication networks). This provides the coexistence of industry with family-based agriculture. In reality, the agricultural structure of these "concelhos" is clearly based on the small farms, which have an average size clearly below the average (2,2 ha against 7 ha).

Besides influencing the shape and the way of occupying the space, this agricultural structure is also one the elements that regulates the organization and the structure of family institution, since it gives the possibility to define family strategies for plural activity and different income sources. We are witnessing an impressive family mobilization towards work, demonstrated by the high activity levels, which is supported by the important women labour participation and abundance of large families. In fact, in average, almost two thirds of "concelhos" have families with more the 3,5 elements (specially concelhos located in the North), which clearly above the national average (3,08). These data supports the idea that family institution has an enormous influence on the local social structure, as also suggestion of Reis (1992, p. 232).

## 4. Conclusion

The proposed methodology has proved to be very appropriate for the identification of the Marshallian industrial districts in Portugal. On the 16 selected "concelhos", we found economical and social structures, typical of local productive systems, defined by Sforzi, in terms of operation (1990) and described by Reis (1992, p. 121) as a socio-territorial entity characterized by the internal interactions of small and medium companies involved in different phases of the same manufacturing process, spatially concentrated, strongly connected to the local population and sharing a limited geographic area.

[9] The cases of Marinha Grande and São João da Madeira, with urbanizing rates of 100%, are representative of "*concelhos*" with only two villages (on the first case) or with only one (on the second case).

## References


[1] Becattini, G. (1979), "Dal'settore industriale al "distretto industrale". Alcune considerazioni sull'unità di indagine dell'economia industriale", *Rivista di Economia e Politica Industriali* 1*:* Jan. 35-48.

[2] Becattini, G. (1990), "The Marshallian industrial districts as a socio-economic notion", in F. Pyke, G. Becattini e W. Sengenberger (eds.), *Industrial Districts and Inter-Firm Co-Operation in Italy*. Geneva: International Institute for Labour Studies.

[3] Beguin, H., 1979, *Méthodes d'Analyse Géographique Quantitative.* Paris: LITEC.

[4] Bianchi, G. (1998), "Requiem for the Third Italy? Rise and fall of a too successful concept", *Entrepreneurship & Regional Development* 10: 93-116.

[5] Brandão, A., Pires, A. e Portugal, J. (1998), "Agrupamentos de concelhos de Portugal Continental e sua caracterização", *Revista de Estatística 1:* 71-94.

[6] Cerejeira da Silva, J.C. (1999), *Distritos Industriais em Portugal Identificação e Avaliação das Externalidades Dinâmicas Associadas,* Dissertação de Mestrado. Braga: Universidade do Minho.

[7] Cliff, A., Haggett, Ord, J., Bassett e Davies (1975), *Elements of Spatial Structure* Cambridge: University Press.

[8] Cooper, L. (1963), "Location-allocation problems", *Operations Research* 11: 331-343.

[10] Everitt, B. (1980). *Cluster Arnalysis.* New York: Halsted.

[11] Garofoli, G. (1983), "Le aree sistema", *Politica ed Economia*, XIV*:* 57-60.

[12] Garofoli, G. (1994), "Os sistemas de pequenas empresas: Um caso paradigmático de desenvolvimento endógeno", in G. Benko e A. Lipietz (eds.), *As Regiões Ganhadoras.* Oeiras: Celta Editora.

[13] Gomes, P., Bacelar, S. e Saleiro, E. (1994), "Contributo para a definição de uma tipologia sócio-económica dos concelhos da Região Norte", *Revista de Estatisticas e Estudos Regionais do INE,* 5: 6-15.

[14] Leandro, A., Esteves, C., e Castro, E. (1997), "O estudo dos sistemas produtivos locais: Reagrupamento dos códigos da CAE segundo o conceito de *cluster*", *Estudos Regionais,* Direcção Regional do Centro: 31-48.

[15] Milligan, G. (1980), "An examination of the efect of six types of error perturbation of fifteen clustering algorithms", *Psychometrica* 45: 325-342.

[16] MacQueen, J. (1967), "Some methods for classification and analysis of multivariate observations", *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability* 1: 281-297.

[17] Murray, A. (1998), "Assessing clustering methods for exploratory spatial data analysis", *38th European Congress of the Regional Science Association,* (Vienna: 26-30 Agosto).

[18] Murray, A., e Estivill-Castro, V. (1998), "Cluster discovery techniques for exploratory spatial data analysis", *International Journal of Geographical Information Science* 12.

[19] Reis, E. (1997), *Estatística Multivariada Aplicada.* Lisboa: Edições Sílabo.

[20] Reis, J. (1985), "Modos de industrialização, força de trabalho e pequena agricultura - para uma análise da articulação entre a acumulação e a reprodução", *Revista Crítica de Ciências Sociais* 15: 16-117.

[21] Reis, J. (1992), *Os Espaços da Indústria: a Regulação Económica e o Desenvolvimento Local em Portugal.* Porto: Edições Afrontamento.

[22] Sforzi, F. (l987), "L'identificazione spaziale" in G. Becattini (ed.), *Mercato e Forze Locali: Il Distretto Industriale.* Bologna: Il Mulino.

[23] Sforzi, F. (1990), "The quantitative importance of Marshallian industrial districts in the Italian economy", in F. Pyke, G. Becattini e W. Sengenberger (eds.), *Industrial Districts and Inter Firm Co-Operation in Italy.* Geneva: International Institute for Labour Studies.

[24] Silva, M.R. (1988), "Developpement industriel et espaces productifs locaux: etude sur la Valée de l`Ave (Portugal)", *Investigação - Economia* 13, Faculdade de Economia do Porto.

[25] Silva, M. R. e Figueiredo, O. (1992), "Analyse comparative de deux systèmes productifs locaux dans le Nord Littoral du Portugal", comunicação apresentada ao seminário internacional *Organisation de la Production et Territoire: Les Systèmes de Petits et Moyennes Enterprises,* Pavia, Itália, Julho.

[26] Sokal, R., e Sneath, P. (1963), *Principles of Numerical Taxonomy.* São Francisco: W.H. Freeman.

[27] Wise, S., Haining, R. e Ma, J. (1997), "Regionalisation tools for the exploratory spatial analysys of health data", in M. Fischer e A. Getis, (eds.), *Recent Developments in Spatial Ana1ysis - Spatial Statistics, Behavioural Modelling and Computational Intelligence.* Berlin: Springer-Verlag.