Innovation in Europe -Classification of European Regions

A multivariate approach using patent data for the chemical industry

Amèrica Grau*

Graduate Student Universitat Pompeu Fabra

america.grau01@aaa.upf.es

Barcelona, November 2002 (revised April 2003)

Abstract

This paper researches on the cluster formation in European regions NUTS 2. We use the number of patents applied at the European Patent Office in the period 1984-1997, as well as the main socioeconomic characteristics of each region. Our goal is to understand which of the main variables included in the model have the most discriminating power and which characteristics show the groups of regions formed. To see why we use patent data we introduce several issues raised in the economic literature relating to the interlink ages between number of patents and innovative capacity of the regions. In addition, it explains the economic rationale for R&D policies and the direction the European Union as a whole is taking in this matter.

* I would like to thank Professor Walter García-Fontes from Universitat Pompeu Fabra for providing me with a dataset on patent data for the European chemical industry, as well as his patience concerning my efforts to advance in this background paper while my internship with The World Bank Group in Washington, DC.

Table of Contents

List of Figures

1 Introduction

Europe is a large economic area with a significant heterogeneity among its regions. This heterogeneity, which at some extent is derived from the different cultures and history existing in the continent, comes mainly in the form of different levels of income per capita, unemployment rate and patterns of specialization, among others. One of them, the concentration of innovative activities, will be the focus of this essay.

In the first part of this essay we will briefly discuss the main issues related to innovation and the rationale of the existing R&D policies within the EU to give us a broader perspective of the simple analysis we will perform in the second section. In addition, we will introduce some topics about innovation in the chemical industry and the role of patents in the innovation process.

In the second section of the essay, we will perform a multivariate analysis using the number of patents applied at the European Patent Office for the period 1984-1997 for each region.[1] We will match this data, together with the level of R&D expenditure in relation to the region GDP, to some socioeconomic characteristics of the region. Our main purpose is to form homogeneous groups using cluster and discriminant analysis, as well as explain which have been the main variables contributing most to the grouping of the regions. We discuss the results and the groups formed at the end of the section. Finally, we present the conclusions of the essay in section 4.

[1] We have classified our regions according to the NUTS2 Classification. However, given that the NUTS2 classification is quite heterogeneous in terms of population size, in some cases we have grouped some NUTS2 regions to obtain a more homogeneous classification in terms of population. Some countries, such as the UK, have not been included in the final data set due to the lack of socioeconomic variables.

2 Background

2.1 Why do we use patent data?

The patent institution was established by the medieval Venetian state, which articulated the basic features of the law today: spur innovation through the incentive of limited-time exclusivity by demanding the demonstration to the public of a working model and promising to seize and destroy counterfeit products. Patent rights arise because inventing is an expensive process and costs must be recouped to provide incentives to invest. If others can cheaply appropriate an inventor's innovation, calling it their own without having invested time and energy in it, investments in innovations will not be made. The patent institution gradually spread northward as Venetian artisans looked for better markets. Patents were granted in France, Germany, and England beginning in the sixteenth century.

Patents matter most to competition in industry sectors characterized by high costs of research, development, and innovation but relatively low costs of imitation. Examples of those industrial sectors can be the research-based pharmaceutical companies and the manufacturers of fine chemicals.

Although not all new inventions are patented and the quality of inventions can vary greatly, patent data is considered a good indicator of inventive activity, together with R&D expenditures. In our analysis we will only use a patent data count in the chemical industry between the years 1986-1992, as well as the percentage of R&D expenditure with respect to GDP, together with a couple of indicators about the socioeconomic characteristics of the region the patent has been granted. However, as Griliches (1990) notes, there is a lot more information that can be derived from patent document than just their aggregated number. For instance, he mentions that things such as the inventions to which the patent is related can be used to identify networks.

2.1.1 Innovation, economic growth and the role of patents

Traditionally, economic growth has been considered the result of the accumulation of two basic production factors: physical capital and labor. The empirical application of the neoclassical model allowed quantifying the contribution of these two factors. It showed that the main engine of growth did not reside on those but on an additional factor that encompassed everything that the model did not take into account. That term was referred as the Solow Residual and it is a measure of the technical progress of an economy. Later, studies appeared which tried to reduce the residual by taking into account new production factors such as innovation, human capital, infrastructures, etc.

Recently, a few economists have tried to explain the residual, the technological progress, because of the decisions taken by the agents in the economy. The decisions taken about the amount of training of workers within a company and the resources spent for research can explain why progress happens in some economic sectors. At the same time, the diffusion of the new inventions, either between sectors or among different economies, allow progress to manifest itself in an aggregate manner and not only within individual companies. The existence of these externalities would allow overcoming the decreasing returns of the production factors and the continuous growth of an economy.

Under the traditional framework, free markets tend to under produce innovation because of an "appropriability problem", thus the government intervenes into the market to provide a period of exclusive distribution right as an incentive to invest in innovation. However, government demands of the innovator a *quid pro quo* for the granting of a patent: As a matter of business strategy, inventors must not safeguard the invention through the darkness of trade secret protection, but rather must disclose how the invention works and in return receive a limited period of exclusive control to produce and license for production. Furthermore, in order to promote the public's interest in the prevention of monopolies from arising through the patent institution, government policy-makers attempt to define concepts of inventiveness, scope, and duration of patent to strike a careful balance between exclusivity and competition as incentives to innovate and distribute. The patent institution aims at encouraging the introduction of new products and processes into the marketplace to benefit the public and propel economic growth. It is an intervention by the government to correct a deficiency of unregulated markets, which if left to them tend to under produce innovation.[2]

In the most prevalent model in the literature of technological change, the knowledge production function, formalized by Griliches (1979), the firm exists exogenously and then engages in the pursuit of new knowledge as an input into the process of generating innovative activity, considering R&D the most important source of new knowledge. Griliches (1989) comments that patents may be a good indicator of the unobserved inventive output and that patent statistics would provide a measure of the innovative activity within a firm (or a region). However, not all inventions are patentable and the quality of the inventions differs greatly. The relationship between R&D investment and patents has long been studied. Griliches (1990) summarizes the literature on the issue and concludes that there is a strong relationship between patent numbers and R&D expenditures. The estimated total elasticity of patents with respect to R&D expenditures lies between 0,3 and 0,6.

Empirically, several studies have shown that R&D activities tend to concentrate geographically and that investment in R&D spills over. The theory of localization suggests that because geographic proximity is needed to transmit knowledge and especially tacit knowledge, knowledge spillovers tend to be localized within geographic regions, something that has been supported in studies such as Jaffe *et al.* (1993) and Audretsch *et al.* (1996). Jaffe *et al.* (1993) compare the geographic distribution of the patent citation with the distribution of the patents cited and find that citations tend to come from closer geographic areas. Additionally, Audretsch *et al.* (1996) find that innovative activity tends to cluster more in industries where

---

[2] In section 2.2 (page 4) we will give a broader perspective of the rational for public intervention.

knowledge spills over play a decisive role. However, the mechanisms that operate for the knowledge to spill over are still not fully clarified.

2.2 Research and Development Policies in Europe

As mentioned above, the regions within the European Union present a great heterogeneity. There are many discussions as to whether the gaps existing between the regions in Europe are closing and the poorer regions are converging towards the economic levels of the richest. However, the evidence seems to show that the convergence processes in terms of income per capita between countries is greater than between regions.[3] One of the factors determining the convergence at the national and regional level is the technological research, development, and innovation.

Given the market failures present in the production and dissemination of knowledge, the European Union as well as each Member State has a set of policies aimed at improving the research and development capacity of the continent. However, those policies present the always-existing trade-off between economic growth and cohesion –that is, the trade-off between equity and efficiency.

Traditionally, those policies have assumed that knowledge could be reduced to information and easily be codified; however, as we mention in the previous section recent development in economics have stressed out that knowledge is not only composed of information, but on tacit knowledge, that is, know-how, skills, etc. The latter cannot be easily transmitted and there exist learning costs in its acquisition. That implies a shift in emphasis of policy, from focusing on incentives to innovate to a focus on diffusion of knowledge (OECD, 1996). Thus, under this context, the main idea of the new policies is to create and environment favorable to innovation and knowledge diffusion; a strong emphasis on network and, in particular, between the relationships between firms and universities and other knowledge institutions.[4]

The fundamental policy problem regarding innovation is related to the presence of market failures in the production and dissemination of knowledge. This prevents Pareto efficient allocation of R&D resources and undermines the private incentives to invest in R&D.[5] The main market failures relevant to the provision of R&D are the externalities, uncertainty and indivisibilities inherent in R&D activities.

---

[3] According to the Second Cohesion Report (European Commission, 2001), between 1988 and 1999 the cohesion countries (Spain, Greece, Portugal and Ireland) there has been a convergence of one-third, whereas of one-sixth between regions. However, it is not the purpose of this essay to discuss the issue and the academic literature referring to it.

[4] For an empirical analysis and evaluation research networks see Garcia-Fontes and Geuna (1999), Garcia-Fontes and Gambardella (1996). Also, see Gambardella and Malerba (1999) for a comprehensive study on the organization of the innovative activity in Europe.

[5] That discussion is based on the PRISM seminar held at the Copenhagen Business School, September 2002, as well as from Stoneman, P. (1995).

Three ways can be considered by the private agents to mitigate the problem with externalities. First, they can try to make knowledge difficult to imitate; second, they can try to internalize the externality by making the new knowledge available in industrial associations, signing R&D collaborative agreements, etc. That solution would allow the firm to take all the returns from the innovative knowledge. However, it is quite costly because of coordinating costs among members of the group. Third, they can try to be a first mover in the applications of its new knowledge and gain a first-mover advantage. Although these actions would be effective in protecting the innovation from appropriation, they hinder the diffusion of the new knowledge to the economy, something that has shown to be favorable to economic growth.

The traditional policy solutions to the market failures have considered the establishment of a market for knowledge, especially with the allowance of property rights and patents. Also, the subsidization of R&D activities; producing innovative activities within public research centers, which then will distribute the new knowledge to firms; and, finally, encouraging the R&D collaboration, an instrument that has been widely used by the European Union.

2.2.1 Main Developments in European Innovation Policies

Most of the innovation policies from the European Union are encompassed under the Framework Programs, which define the priorities in terms of research and development, are elaborated for five years, and are justified under the grounds of market failure arguments mentioned in the previous section. The objective of the Framework Programs is to favor cross border collaboration to exploit spillovers; as well as to build a research and innovation equivalent of the common market for goods and services. They provide for the possibility to exempt aid to R&D from the general prohibition of state aids.[6] Broadly speaking, the Framework distinguished three levels of aid intensity –that is, aid as a percentage of the cost of R&D. First, basic research -100% funding, second, industrial research -50% of funding, and third pre-competitive development activity, with 25% of funding.[7]

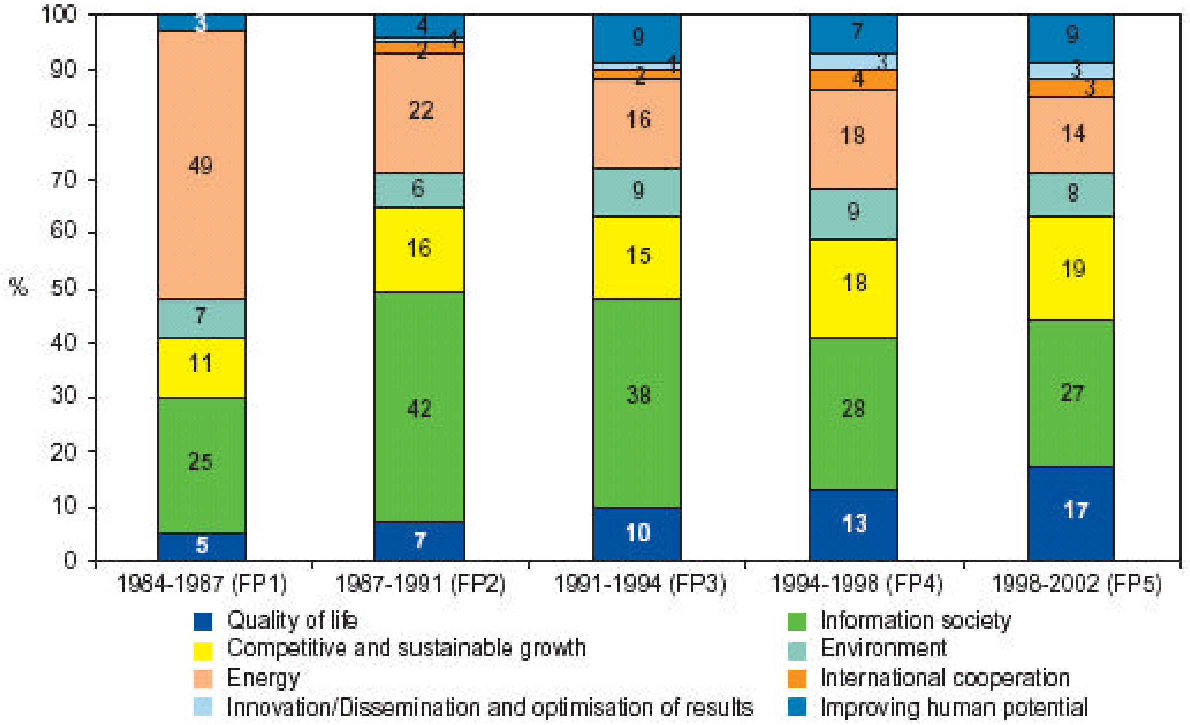
The First Framework Program 1984-87 concentrated on the need to develop alternative energy sources and allocated 50 percent of its funds to energy R&D. The following Framework Programs concentrated on stimulating advances in information technology (ESPRIT) and communications technology (RACE) and followed a more integrated strategy towards innovation. Because of the Green Paper of Innovation published in 1995 (European Commission, 1995), which identified one of Europe's major weaknesses its difficulty to translate the results of technological research and skills into innovations and competitive advantage, the First Action Plan on Innovation was created (European Commission, 1996). It called for establishing an innovation culture in Europe, a legal, regulatory and financial framework conductive to

[6] Some argue that the current level of subsidies under the Framework Program is for, example in the case of the aeronautical research, a backdoor subsidy, thus affecting competition. http://www.economist.com/PrinterFriendly.cfm?Story_ID=1324671

[7] There is some literature on the effectiveness of public incentives to R&D. For a review, see Bronwyn, H., Reenen, J.V. (2000): "How effective are fiscal incentives for R&D? A review of the evidence." Research Policy, April, Vol. 29, Iss. 4/5.

innovation and to gear research towards, by that meaning intensifying the research links between universities and companies and a stronger promotion of start-ups.

Figure 1 Evolution of Priorities of RTD Framework Programmes



Source: European Commission (2000a)

The Fifth Framework Program (5FP) 1999-2002 concentrated on four thematic programmes and three horizontal programmes with an overall budget of ECU 14.960. It consisted of a limited number of research areas, which combine technological, industrial, economic, social and cultural aspects. In response to the First Action Plan, it introduced the concept of "key actions" that mobilized a wide range of scientific and technological disciplines that are required to approach one specific problem, instead of mobilizing a specific discipline. It has two distinct parts, the European Community framework program covering research, technological development and demonstration activities; as well as the Euratom framework programmed that covers research and training aspects in the nuclear sector.

The Sixth Framework Program 2002-2006, which is about to start as this essay is being written, has one of its main priorities the compromise reached in the Lisbon Summit 2000 to "transform the European Union into the most dynamic and competitive knowledge-based economy in the world" (sic).[8] The budget contribution with respect to the previous Framework Program is of 17% and making up 3,9% of

---

[8] See European Commission, 2002 pp. 21-22 for the Commission recommendation on how to attain this objective by improving the R&D and innovation performance.

the Union's total budget. The seven key areas chosen in the program have been genomics and biotechnology for health; information society technologies; nanotechnologies and nanosciences; aeronautics and space; food safety; sustainable development; and economic and social sciences.

Since January 2000, the Trend Chart on Innovation tracks the major innovation policy developments in all EU Member States, plus some additional countries, according to the foundations lay out by the First Action Plan. It reports a cross-country comparative analysis with the aim of highlighting the major trends and developments in national innovation policies by each Member State. With this information, an interesting database of the measures presented by each country was built, specifying the action plan in which they were included as well as the dates in which it operated.[9] Its main findings were the given emphasis of all Member States to the promotion of clustering and co-operation of innovation and gearing research to innovation, as well as the development of a strategic vision of R&D. In contrast, issues of creating and maintaining an innovation culture, establishing a framework conductive to innovation and the promotion of competition were not equally targeted among countries.

A piece that still needs to be solved within the European R&D policy is the establishment of a Community Patent. In Europe, patents can be filed on a national basis, through the patent offices of individual European countries. In addition, there exists a European Patent Office, which grants patents according to the European Patent Convention, but is not exclusive to EU Member States. Under the European Patent Convention, patent protection is achieved through a single application resulting in individual national patents in the countries where they are desired, but does not give raise to protection in all EPC countries automatically. And not only that, but there system does not provide for a court with powers to settle patent disputes at European level; the risk being that the competent courts in the Member States can hand down contradictory judgments about a case.

A Community Patent giving automatic patent protection for all EU Member States through a single application is an idea that dates back to the 1960s. However, there are concerns that the costs might be too high due to translation requirements and the industry is worried about the effectiveness of its proposed system for revocation.

2.3 Innovation in the Chemical Industry in Europe

The chemical industry in Europe is the third largest manufacturing sector and it accounts around 10% of EU manufacturing industry's gross value added (CEFIC, 2001). It produces a wide range of variety of products subject to different R&D requirements. Because of the high capital required and the economies of scale, basic chemicals tend to be produced by large firms, whereas fine and specialized chemicals, whose R&D activities are focused on process and product development and developing new products, tend to be produced by smaller firms.

The R&D activity of the chemical industry tends to be higher than the industrial average, although it seems that there exist diminishing returns to R&D, the exceptions being the pharmaceutical and biotechnology field. Results from Albach *et al.*, (1996) show that the most important objective of innovation is to increase the efforts to applying strategies of cost leadership to be able to compete in price. In addition, those strategies aimed at reducing cost to compete based on price were found to be important drivers of innovative activity, as well as the specialization in certain product areas to acquire a higher market share in certain products. The same study shows that the most innovative sub sectors within the chemical industry between the periods 1984-1993 are chemicals, paints/varnishes and plastics. The major factors impeding innovation were the financial risk associated with the innovative activity as well as difficulties to access capital.

Innovation in any industry can take two forms: process innovation and product innovation. Traditionally, product innovation has been considered a strategy followed by firms to gain competitive advantage. However, process innovation, which accompanies product innovation, is an element that enters in the strategic equation of chemical firms because it allows offering a cheaper product to the customer –which can also be referred as cost leadership.

The choice between one and another form of innovation is explored in several theoretical studies, as well as the existence of cooperative R&D activities with a firm's potential rivals. For example, Rosenkranz (1995) shows that welfare is increased if firms coordinate their research activities and share R&D costs. When the firms cooperate but their R&D costs are not shared, she shows that welfare is only enhanced if product innovations are not too expensive.

In Albach *et al.* (1996) they use a logit regression model to estimate the likelihood of a firm being either a product innovator or a process innovator in the chemical industry in Europe for the period 1984-1993. Their results show that a process innovator shows a high degree of commitment to the economical use of resources and of competing with rivals. Additionally, they also find tat when product innovations are easy to copy, the probability of being a process innovator decreases. The probability of being a product innovator increased both by the commitment with innovating products and an awareness of the risks involved in the innovative process.

From an economic point of view, firm incentives' to innovate are the sunk cost effect, the replacement effect and the efficiency effect. The sunk cost effect relates to the investment committed by a firm that has already introduced an innovation and a firm that hasn't. The costs associated with that investment are sunk and should not be taken into account when deciding about making innovation. However, the sunk cost tights the firm to the current innovation. On the other hand, the other firm considering entering the market faces a different situation since it can compare the cost of every alternative innovation that is considering. In the case of the chemical industry, a considerable share of its R&D expenditures can be

---

[9] See pages 109-132 of the report (European Commission, 2000b).

considered sunk costs, although some of the knowledge acquired in developing a specific product or improving a process can be used for future undertakings (Fleischer *et al.*, 2000).

The second incentive has to do with the replacement effect. Following Fleischer *et al.*, we will use Arrow's concept of replacement effect of any innovation process to a product innovation. The former tends to have a lower average variable cost. Arrow compared two situations, a situation with a monopolist using an old technology, and another situation with a potential entrant that might become a monopolist. If both companies have equal innovative capabilities, an entrant would be willing to spend more than the monopolist would because that would give him access to become a monopoly, whereas the monopoly would confirm his position. That explains why because of the dynamic of the market an innovative entrant may overtake an established firm.

The efficiency effect has to do with the incentives that the monopolist has to innovate due to the threat of entrance of an innovator and the losses that this would entail in terms of both business and price competition for the monopolist.

Which effect dominates depends on the conditions of the innovation competition. Fleishcer *et al.* (2000) suggest that in the European chemical industry the replacement and sunk cost effect may dominate, since the chances a small company will enter the markets are low. However, they point out that the current changes in the competitive environment of the chemical industry might be increasing the efficiency effect.

2.3.1 Patents in the Chemical Industry

Research-based chemical and pharmaceutical companies are the most dependent on the patent as a matter of competitive advantage, for their chemical products can be easily pirated by unscrupulous competitors unburdened by the investment in research and development.

The chemical industry's dependence on patents is rooted in its history and structure. At least since the mid-nineteenth century it has depended on innovative products: coal tar dyes, explosives, and synthetic fibers. In turn, the English, French, German, and American industries thrived. These competitors were growing within an environment that was offering good patent protection, adequate financing, sophisticated chemical science education, and capable management. During the 1920s, mergers, acquisitions, and failures resulted in market consolidation in the United States, Germany, Great Britain, and Switzerland. Until the 1974 oil shock, the chemical industry was the world's largest and fastest growing industrial sector.

Investment demands are high in the basic commodity chemicals and the fine, high technology chemicals. Yet, competitive advantage in commodity chemicals rests on manufacturing scale economies; competitive advantage in high-technology chemical rests on proprietary product and process technology.

The world's largest chemical makers are based in Europe: Bayer, Hoechst, BASF, and ICI. Europe offers certain major advantages and disadvantages for the chemical industry. Disadvantages include high costs of capital, energy and manpower, and the costs associated with complying with stringent health, safety and regulations. One of Europe's main advantages, however, is the strength of its science and engineering base, which is a rich source of world class research expertise and highly trained people: both of these resources are vital to the chemical industry, and indeed any technology-based industry.

The chemical industry is heavily dependent on research and development to deliver new products and processes, improving environmental performance, enhancing the quality of life and sustaining a competitive edge. An effective patent legislation encourages companies to innovate and the tools for preventing the piracy of innovations; it rewards innovators, providing the incentives for a sustained investment in research and development and, finally, it provides a valuable source of research and technical information for industry and the science and engineering base, thus promoting the development of technology.

Patents are a driving force to provide the means necessary for the cycle of investment, research, innovation and re-investment. Europe's attraction to international investment is dependent on its ability to provide a legal framework that stimulates, facilitates and rewards industrial innovation. The current pattern legislation and the lack of an integrated Community Patent pose a serious barrier to the exploitation of this market by high technology business, and the chemical industry in particular, patents being a very important indicator of industrial performance in this industry (Achilladelis and Schwarzkopf, 1990).

Figure 2 Share of European (EPO) Patents by Field of Technology, %

	Sare of total European patents 1997	Average annual growth in European Pattents 1991-1997
Electrical engineering		
Electrical machinery and apparatus, electrical energy	7,0	-0,1
Audio-visual technologies	2,8	-0,1
Telecommunications	5,6	6,9
Information technologies	2,7	9,4
Semiconductors	1,3	1,8
Instruments		
Optics	2,3	-1,0
Analysis, measurement, control technology	7,0	-0,5
Medical technology	3,9	6,5
Nuclear engineering	0,4	-7,7
Chemistry, pharmaceuticals		
Organic fine chemistry	3,5	-6,0
Macromolecular chemistry, polymers	3,3	-3,4
Pharmaceuticals, cosmetics	5,0	20,9
Biotechnology	1,5	0,8
Chemical and petrol industry, basic materials chemistry	2,9	4,0
Process engineering, special equipment		
Materials, metallurgy	1,9	-1,3
Handling, printing	6,1	0,7
Environmental technology	0,8	0,7
Mechanical engineering, machinery		
Machine tools	3,2	2,3
Engines, pumps, turbines	2,6	2,4
Transport	6,4	6,1
Space technology, weapons	0,6	-4,5
Consumer goods and equipment	10,3	2,7
Other	19,1	1,6
All technologies	100,0	2,0

Source: European Commission (2000a)

3 Multivariate Analysis

In this section of the paper, we will perform a simple cluster analysis for the European regions using data on patents registered at the European Patent Office under the chemical section as well as its GDP per capita and population. We will use four variables in our analysis:[10] the number of patents filed (PATENT), the percentage in research and development expenditure with respect to GDP (RDEXPPER), the income per capita with respect to the European average (GDPPC) and the population size (POPUL).[11]

We have data for the regions within the Member States of Austria, Belgium, Germany, Denmark, Spain, Finland, France, Greece, Italy, the Netherlands, Portugal and Sweden.

Our aim is to see which have been the variables contributing most to the grouping of regions and which groups of regions are formed. We must keep in mind that cluster analysis is a pure descriptive tool; no inference can be made of the results we will obtain.

3.1 Methodology

Our clustering exercise attempts to classify our regions in groups such that the regions in each group display similar features because the technological composition of the region of main patent holders is similar.

1. Cluster Analysis. Cluster analysis is a pure descriptive tool that locates the elements of the sample (in our case the regions) in different groups. There are many methods to establish a classification; the grouping of regions will differ according to the method used. The key to cluster analysis is to recognize when these groups formed are real or whether they are mere impositions of the method. Consequently, we need to statistically validate the result obtained using a discriminant analysis.

   We have used a method of classification in two phases: first, we have used a hierarchical classification (it does not need a previous definition of the number of groups) to determine the ideal number of groups. Then, we have used a non-hierarchical classification (the researcher must include the number of groups) based on the information provided by the first method.

   To study the similarities between the regions we have used the squared Euclidean distance

[10] We would have liked to include more technological and socioeconomic variables in the analysis; with those, we would have constructed some fictitious variables using factor analysis techniques prior to performing the cluster and discriminant analysis. However, we were unable to easily find available data for all regions. Because of this, we have restricted the data set. In addition, in the cluster analysis we have only included the least correlated variables of all the variables we had.

[11] Period covered for each variable: RDEXPPER 1992, except Greece 1993; PATENTS 1984-1997; GDPPC 1992, except FR9 1994; POPUL 1992, except DED 1994.

method. The squared Euclidean distance is sensitive to the scale used in the variables; for that reason, the variables need to be standardized so that the bias is eliminated.

The hierarchical method of classification begins by assigning a group to each case. In the first step, the squared Euclidean distance determines the distance between cases. Later, when it comes to grouping the regions, we need to specify which aggregation method will be used. In our case, we will use the Ward's method. This method uses an approximation to the Analysis of Variance to evaluate the distance between clusters, minimizing the residual sum of squares of the two clusters formed in each step.

The non-hierarchical method (K-means) is used based on the information provided by the dendogram resulting from the hierarchical method. It can be considered as a backwards Analysis of Variance, meaning that the significance test for the null hypothesis that the means of the groups are different in ANOVA evaluates the variability between groups along the variability within groups. This classification method moves the elements or cases from a group to another until the most significant ANOVA results are achieved.

2. Canonical discriminant analysis to evaluate and explain the classification. It evaluates the goodness of fit and explains the relevant variables in the discrimination between groups. It determines the contribution of each classification variable for each region, as well as the determination of the group to which a region belongs. Thus, it will be possible to determine whether a region has been correctly classified or not by the cluster analysis. It is a similar technique to that of the regression analysis, except that the dependent variable is a categorical variable. The criterion used to obtain the discriminant function is the maximization of the within and between groups variability ratio.

3.2 Preliminary Exploration of the Variables

As can be seen in Figure 3, the regions presenting the highest patent applications are Ïlle de France and Nordrhein-Westfalen, accounting for almost the 20% of the patent applications in the chemical industry for the period considered. They also show a higher than the average income per capita. On the other hand, the regions with the lowest number of patent applications are located in Spain (Extremadura and La Rioja).

The number of patents in each region presents a great variability. The descriptive Statistics of the Variables (see Figure 4) show us that the regions are quite heterogeneous in terms of population and number of patents assigned. The lowest variability can be found in the R&D expenditure. In order to avoid any bias that might occur due to the different scale measures when performing the cluster analysis, we will standardize all variables of the matrix of data.

It will also be useful to calculate the Pearson Correlation coefficient among all variables. The Pearson correlation coefficient is a measure of lineal relationship between variables. The correlation coefficient can lie between –1 (a perfectly negative relationship) and +1 (a perfectly positive relationship). A value of zero indicates that there is no lineal relationship between the variables. When interpreting the results, we have to keep in mind that no causality can be derived out of a significant correlation.

Figure 5 shows that the correlation coefficients between the number of patents and the population number are quite high (0,714). The rest of correlation coefficients are relatively low, with the income per capita and the population size having the lowest one (0,244).

3.3 Results of the Analysis

3.3.1 Cluster Analysis

Cluster analysis will allow us to classify a great amount of information in smaller groups that have a meaning. Since the goal of the analysis is to create relatively homogeneous groups, in the dendogram resulting from the analysis we will have to decide how many groups to form. We have to "cut" the graph in such a way that we obtain the optimal number of groups. On the other hand, in a non-hierarchical cluster the researcher has to specify the number of groups in the data, something that requires an iterative and trial-and-error method.

Two potential outliers can be identified in the dendogram (a branch that unifies very late). In our case, those potential outliers are the regions of DEA and FR1 (Nordrhein-Westfalen and Île de France), the same regions we identified in the previous section as having a very significant proportion of patent applications.

Figure 7 shows that the biggest percentage change occurs between cluster one and two, followed by from two to three, and from five to six. We will consider the formation of three and five clusters.

Using these results, we have performed a non-hierarchical cluster analysis, both for the case of three clusters and for the case of five clusters. The classification results can be seen in Figure 8 and Figure 9.

3.3.2 Discriminant Analysis

In order to validate the results presented above we need to perform a discriminant analysis, using the cluster numbers assigned to each region as the categorical variable. Since we have worked with the formation of both three and five clusters, we will perform the discriminant analysis for those two cases and see when we achieve a better grouping.

An important statistic we will analyze is the Wilk's Lambda. The Wilk's Lambda is a statistic computed to show the discriminating power of the model. Its value is comprised between zero and one. A value of lambda equal to zero means a perfectly discriminating power. In addition, the Wilk's Lambda can be transformed into an F-statistic.

The discriminant analysis generates a discriminant function based in the lineal combination of the variables included in the analysis that best discriminate among groups. For each function, we will obtain statistics such as the Wilk's Lambda and the eigenvalues. The eigenvalue of a function can be interpreted as the share of total variability of the scatter dots projected through the whole range of functions of the model. If the value is high, the function will have a high discriminatory power. If the value of the canonical correlation is high (close to 1), the dispersion will be caused by the difference between groups, and then the function will have a high discriminatory power.

In addition, we will compute the coefficient of the standardized discriminant functions generated. Those coefficients will tell us which variables are contributing most to the discrimination of each function. Finally, we will represent the clustered regions in a two dimensional pane where each axes will correspond to a discriminant function.

1. Three Clusters

In Figure 10 we can see that the mean of each variable is significantly different in each group (p-value lower than 0,05). Accordingly, the Wilk's Lambda is low for all variables, meaning that each of them has a relatively high contribution to the model. The variable contributing most to the grouping are the percentage of R&D Expenditure (RDEXPPER) and the number of patents (PATENTS).

The p-values of the Wilk's Lambda for each discriminant function shown in Figure 11 certify the significance of the discriminant functions, which means that their explanatory capacity is good (they discriminate groups correctly).

Figure 12 shows the eigenvalues of the canonical discriminant functions, which measure the deviations of the discriminant punctuation between groups in respect to the deviation within groups. In our case, functions one and two have a high canonical correlation and their eigenvalue is high compared to function three. In addition, the cumulative percentage of the variance explained by function one is very high (95.4%); the first function is going to give the majority of the discrimination, something we already saw in the Wilk's Lambda.

Looking at Figure 13, the results from SPSS indicate that the variables we have considered are all included in the analysis. This will allow us to form groups according to different socioeconomic and technological characteristics.

Results of the evaluation of the grouping of regions performed in the previous section are shown in Figure 14. As we can see, 98,7% of the grouping has been done correctly. There is only one misclassified region in Group 3. Therefore, it seems that the grouping computed in the previous section has been quite successful. The misclassified region is Comunidad de Madrid.

In order to classify the three groups we need, at least, two discriminant functions. We have obtained the standardized canonical discriminant function coefficients. Figure 15 shows the weights of each variable in every function. Thus, in function one and two RDEXPPER is the variable contributing most, whereas POPUL is the second contributing most in function one but not in function two.

The graphical representation of the functions can be seen in Figure 16.

Summing up, in both functions RDEXPER are highly discriminative, but the most significant change in function two is the high discriminative power of PATENTS in function 2 with respect to function 1.

2. Five Clusters

We will now analyze our results considering the formation of five clusters. We will see which are the variables considered in the model as well as their weight in the discriminant functions.

Figure 17 shows that the mean of each variable is significantly different in each group. Additionally, the Wilk's Lambda has the lowest value (a higher discriminating power) for the variables number of patents (PATENTS) and income per capita (GDPPC). In that sense, the variable contributing least to the model is the population (POPUL).

The functions generated by the discriminant analysis, which are a lineal combination of the variables included in the analysis, show us a high discriminating power (value of Wilk's Lambda is low).

Figure 19 shows the eigenvalues of the canonical discriminant functions. As we mentioned above, a high eigenvalue means a high discriminatory power, the same can be said about the canonical correlation. Thus, function one is the one that best discriminates among groups, with a percentage of the variance explained of 78,2%. Together with function 2 they explain the 97,8% of the total variance.

The variables considered in the analysis are the number of patents (PATENTS), the income per capita (GDPPC) and the percentage of R&D expenditure (RDEXPPER). That leaves us with the population (POPUL) excluded from the model.

When evaluating the grouping of regions, the number of regions correctly classified is 93.3%, a bit lower than in the case of three clusters. There are five regions misclassified (Figure 22).

In order to classify the five groups we need, at least, four discriminant functions. However, in our case, only the first three discriminant functions have been considered in the analysis. The standardized canonical discriminant function coefficients show the weight of each variable in every function and we can then see which variables are contributing most. In function 1, the variable contributing most is the number of patents (PATENTS), whereas in function to it is the income per capita (GDPPC), and the percentage or R&D expenditure in function 3.

If we only consider function 1 and 2 we can obtain a graphical representation of the clusters along these two dimensions (Figure 24).

We can see that Function 1 discriminates between group 3 and the rest, whereas function 2 discriminates against group 2 and the rest. Group 4 and 5 are relatively close one to the other and concentrate a high number of observations. On the other hand, group three has only two observations.

3.4 Discussion of the Multivariate Analysis

The following table summarizes the main differences between clustering the regions in three and 5 clusters:

	3 clusters				5 clusters			
Variables included in the discriminant analysis	RDEXPPER, PATENTS, GDPPC, POPUL				RDEXPPER, PATENTS, GDPPC			
Variables contributing most in the grouping	RDEXPPER, POPUL				PATENTS, GDPPC			
Discriminant Functions		Function 1	Function 2	Function 3		Function 1	Function 2	Function 3
	RDEXPPER	0,699	0,538	-	POPUL	-	-	-
	POPUL	0,630	-0,389	-	PATENTS	0,754	-0,632	-0,200
	GDPPC	0,449	0,308	-	GDPPC	0,452	0,790	-0,427
	PATENTS	0,338	-0,414	-	RDEXPPER	0,414	0,320	0,857
Classification Results. % correctly classified regions	98,70%				93,30%			

The most remarkable difference between the clustering of regions between three and five clusters is the role the variables PATENT and RDEXPPER take in the discriminant functions. In the case of five clusters, the number of patents is the variable contributing most to function one. However, in the case of three clusters the number of patents is the variable contributing less to the first discriminant function. On the other hand, the percentage of R&D expenditure (RDEXPPER) has the highest coefficient in function 1 of three clusters, but the lowest in function 1 of five clusters.

It is interesting to find that the technological variables have a stronger power in discriminating among regions against the socioeconomic variables that we included –especially in the case of five clusters. We are aware of the limitations of this analysis and the fact that we only use patents for the chemical industry, but the evidence seems to confirm what we said in the first section of the essay about the linkages between innovation and the development of regions. Also, there seems to be a high degree of concentration in that industry showed by the fact that almost 20% of the patents issued belong to just a couple of regions: Ille de France and Nordrhein-Westfalen.

Group 1: Regions ranking second in the number of patents in the chemical industry, and a GDP per capita around 12 points above the average. In addition, they show a relatively high amount of total R&D expenditure with respect to the income per capita. It consists of regions in Austria, Germany, France and Sweden.

Group 2: Regions with a very high income per capita and good levels of R&D spending, but low levels of patent applications in the chemical industry. The fact that those regions have a very high income per capita suggests that they are either specialized in sectors not necessarily linked to the chemical industry, maybe sectors which are also very knowledge intensive by the fact of their high R&D expenditures. It consists of two regions in Germany (Bremen and Hamburg), as well as Brussels.

Group 3: Those regions present the bulk of the patent applications for the period considered in the chemical industry (around 20% of all patents). It consists of just two regions, Ille de France and Nordrhein-Westfalen in Germany, but their population size is greater than the average. They show the highest percentage of R&D expenditures and are the second highest group in terms of income per capita. However, they only account for around 9% of the population considered in the sample.

Group 4: Regions with the lowest level of all the variables considered. Their number of patent applications for the chemical industry is around 9% of the average in Europe, their income per capita is around 70% and their R&D expenditures are extremely low (0,7%). What is most worrying, is that these regions represent around 24% of the population considered in the sample. It consists of regions in southern Spain, southern Italy and Greece.

Group 5: Regions with a degree of patent application higher than group 2, but lower percentage of R&D expenditures and the income per capita that is just around the European average. They consist of regions in most France, Denmark, the Center and North of Italy, some regions in the former Eastern Germany, the Netherlands and three regions of Spain (País Vasco, Comunidad de Madrid and Catalunya).

4 Conclusions

In the first part of the paper, we have discussed the main issues around innovation, its links with economic growth, and the use of patent data as an indicator of the innovative level of an economy/industry. We have also discussed the rationale for research and development policies, especially in a heterogeneous continent such as Europe, as well as the tendency of innovative activities to cluster spatially. We have also discussed the importance of innovation within the chemical industry and the significant amount of patents filed in Europe.

In the second section of the essay, we have used patent data for the chemical industry in Europe to explore the clusters formed in the European regions considered. The results have shown us that, as we should expect from the discussion presented in the first part of the essay, we can form more or less homogenous clusters of regions according to its technological characteristics and patent applications in the chemical industry. We have identified five clusters of regions in which patent data and GDP per capita are the most relevant variables. Of the five groups considered, there is one group (group 3), which accounts for 20% of the total patent application in the chemical industry for the period 1984-1997, Ille de France and Nordrhein-Westfalen in Germany, and a GDP per capita above 50% of the European average. At the other end, there is another group (group 4) with both the number of patent application and the GDP per capita extremely low. These results seem to suggest that there is room to foster R&D policies in Europe to reduce, at some extent, the great disparities among its regions, specially when we seen that those regions with the lowest GDP per capita also show the lowest levels of R&D expenditure.

5 Figures

Figure 3 Regions with the highest % of Patent Applications

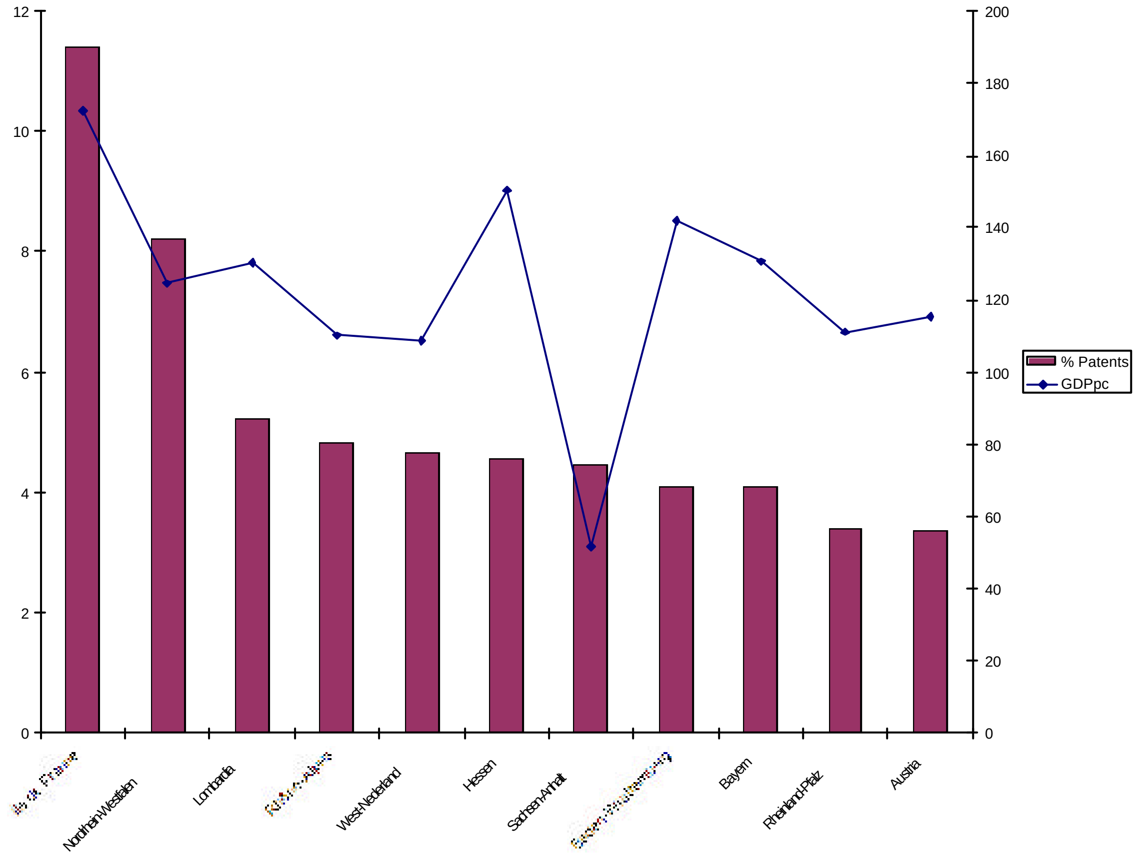


Figure 4 Descriptive Statistics of Variables

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PATENTS	75	1,00	5751,00	671,6933	1015,27853
RDEXPPER	75	,30	3,60	1,3839	,88314
GDPPC	75	45,65	215,70	97,5263	31,86129
POPUL	75	260300,00	1,8E+07	3907023	3091621,414
Valid N (listwise)	75				

Figure 5 Correlation Matrix

(Pearson Correlation Coefficient)

Correlations

		PATENTS	RDEXPPER	GDPPC	POPUL
PATENTS	Pearson Correlation	1	,563**	,441**	,714**
	Sig. (2-tailed)	,	,000	,000	,000
	N	75	75	75	75
RDEXPPER	Pearson Correlation	,563**	1	,521**	,368**
	Sig. (2-tailed)	,000	,	,000	,001
	N	75	75	75	75
GDPPC	Pearson Correlation	,441**	,521**	1	,244*
	Sig. (2-tailed)	,000	,000	,	,035
	N	75	75	75	75
POPUL	Pearson Correlation	,714**	,368**	,244*	1
	Sig. (2-tailed)	,000	,001	,035	,
	N	75	75	75	75

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 6 Hierarchical Cluster Analysis Results. Dendogram for 75 cases

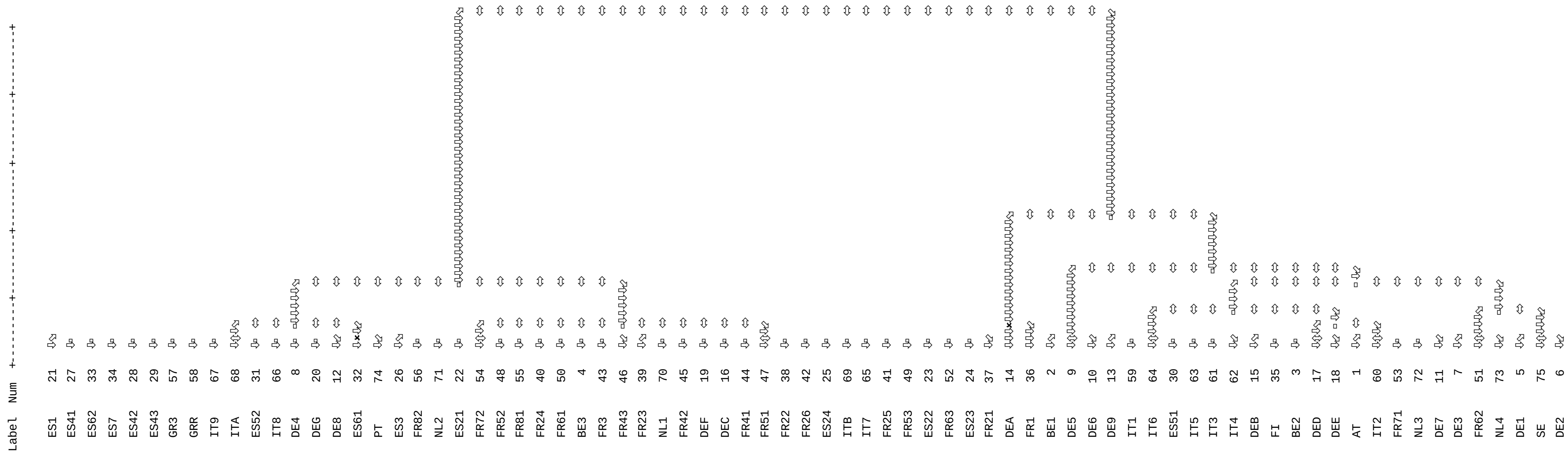


Figure 7 Agglomeration Coefficient Analysis for the hierarchical cluster

Cluster Number	Agglomeration Coefficient	Percentual Change in the coefficient of the next level
10	47,2	13,65
9	53,6	13,03
8	60,6	18,03
7	71,5	17,00
6	83,7	20,98
5	101,3	18,79
4	120,3	20,35
3	144,8	28,68
2	186,3	58,88
1	296,0	-

Figure 8 Non-hierarchical Cluster Analysis Results (3 clusters)

Region	*Code*	*C*	*Region*	*Code*	*C*	*Region*	*Code*	*C*
Baden-Württemberg	DE1	1	Region Wallonne	BE3	3	Basse-Normandie	FR25	3
Bayern	DE2	1	Brandenburg	DE4	3	Bourgogne	FR26	3
Nordrhein-Westfalen	DEA	1	Mecklenburg-Vorpomme	DE8	3	Nord - Pas-de-Calais	FR3	3
Île de France	FR1	1	Saarland	DEC	3	Lorraine	FR41	3
Lombardia	IT2	1	Sachsen-Anhalt	DEE	3	Alsace	FR42	3
West-Nederland	NL3	1	Schleswig-Holstein	DEF	3	Franche-Comté	FR43	3
Sweden	SE	1	Thüringen	DEG	3	Pays de la Loire	FR51	3
Austria	AT	2	Noroeste	ES1	3	Bretagne	FR52	3
Arr. Admin. Bruxelle	BE1	2	País Vasco	ES21	3	Poitou-Charentes	FR53	3
Vlaams Gewest	BE2	2	Comunidad Foral de N	ES22	3	Aquitaine	FR61	3
Berlin	DE3	2	La Rioja	ES23	3	Limousin	FR63	3
Bremen	DE5	2	Aragón	ES24	3	Auvergne	FR72	3
Hamburg	DE6	2	Comunidad de Madrid	ES3	3	Languedoc-Roussillon	FR81	3
Hessen	DE7	2	Castilla y León	ES41	3	Attiki	GR3	3
Niedersachsen	DE9	2	Castilla-la Mancha	ES42	3	Rest of Greece	GRR	3
Rheinland-Pfalz	DEB	2	Extremadura	ES43	3	Nord Est	IT3	3
Sachsen	DED	2	Cataluña	ES51	3	Emilia-Romagna	IT4	3
Finland	FI	2	Comunidad Valenciana	ES52	3	Centro	IT5	3
Midi-Pyrénées	FR62	2	Andalucía	ES61	3	Abruzzo-Molise	IT7	3
Rhône-Alpes	FR71	2	Región de Murcia	ES62	3	Campania	IT8	3
Provence-Alpes-Côte	FR82	2	Canarias	ES7	3	Sud	IT9	3
Nordovest	IT1	2	Champagne-Ardenne	FR21	3	Sicilia	ITA	3
Lazio	IT6	2	Picardie	FR22	3	Sardegna	ITB	3
Oost-Nederland	NL2	2	Haute-Normandie	FR23	3	Noord Nederland	NL1	3
Zuid-Nederland	NL4	2	Centre	FR24	3	Portugal	PT	3

Figure 9 Non-hierarchical Cluster Analysis Results (5 clusters)

Region	Code	C	Region	Code	C	Region	Code	C
Austria	AT	1	Brandenburg	DE4	4	Region Wallonne	BE3	5
Vlaams Gewest	BE2	1	Mecklenburg-Vorpomme	DE8	4	Saarland	DEC	5
Baden-Württemberg	DE1	1	Thüringen	DEG	4	Sachsen-Anhalt	DEE	5
Bayern	DE2	1	Noroeste	ES1	4	Schleswig-Holstein	DEF	5
Berlin	DE3	1	Comunidad Foral de N	ES22	4	País Vasco	ES21	5
Hessen	DE7	1	La Rioja	ES23	4	Comunidad de Madrid	ES3	5
Niedersachsen	DE9	1	Aragón	ES24	4	Cataluña	ES51	5
Rheinland-Pfalz	DEB	1	Castilla y León	ES41	4	Picardie	FR22	5
Sachsen	DED	1	Castilla-la Mancha	ES42	4	Haute-Normandie	FR23	5
Finland	FI	1	Extremadura	ES43	4	Centre	FR24	5
Rhône-Alpes	FR71	1	Comunidad Valenciana	ES52	4	Bourgogne	FR26	5
Lombardia	IT2	1	Andalucía	ES61	4	Nord - Pas-de-Calais	FR3	5
West-Nederland	NL3	1	Región de Murcia	ES62	4	Lorraine	FR41	5
Zuid-Nederland	NL4	1	Canarias	ES7	4	Alsace	FR42	5
Sweden	SE	1	Champagne-Ardenne	FR21	4	Franche-Comté	FR43	5
Arr. Admin. Bruxelle	BE1	2	Basse-Normandie	FR25	4	Pays de la Loire	FR51	5
Bremen	DE5	2	Poitou-Charentes	FR53	4	Bretagne	FR52	5
Hamburg	DE6	2	Limousin	FR63	4	Aquitaine	FR61	5
Nordrhein-Westfalen	DEA	3	Attiki	GR3	4	Midi-Pyrénées	FR62	5
Île de France	FR1	3	Rest of Greece	GRR	4	Auvergne	FR72	5
			Abruzzo-Molise	IT7	4	Languedoc-Roussillon	FR81	5
			Campania	IT8	4	Provence-Alpes-Côte	FR82	5
			Sud	IT9	4	Nordovest	IT1	5
			Sicilia	ITA	4	Nord Est	IT3	5
			Sardegna	ITB	4	Emilia-Romagna	IT4	5
			Portugal	PT	4	Centro	IT5	5
						Lazio	IT6	5
						Noord Nederland	NL1	5
						Oost-Nederland	NL2	5

Figure 10 Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
Zscore(RDEXPPER)	,367	62,123	2	72	,000
Zscore(PATENTS)	,373	60,603	2	72	,000
Zscore(GDPPC)	,610	23,039	2	72	,000
Zscore(POPUL)	,458	42,591	2	72	,000

Figure 11 Wilk's Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	,126	145,968	8	,000
2	,796	16,066	3	,001

Figure 12 Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	5,313[a]	95,4	95,4	,917
2	,256[a]	4,6	100,0	,451

a. First 2 canonical discriminant functions were used in the analysis.

Figure 13 Variables included in the analysis

Step		Tolerance	F to Remove	Wilks' Lambda
1	Zscore(RDEXPPER)	1,000	62,123	
2	Zscore(RDEXPPER)	,926	54,833	,458
	Zscore(POPUL)	,926	36,851	,367
3	Zscore(RDEXPPER)	,926	33,723	,280
	Zscore(POPUL)	,874	40,274	,306
	Zscore(GDPPC)	,941	9,247	,180
4	Zscore(RDEXPPER)	,916	26,162	,222
	Zscore(POPUL)	,763	13,289	,175
	Zscore(GDPPC)	,932	7,366	,153
	Zscore(PATENTS)	,873	4,450	,142

Figure 14 Classification results

			Predicted Group Membership			
		Cluster Number of Case	1	2	3	Total
Original	Count	1	7	0	0	7
		2	0	18	0	18
		3	0	1	49	50
	%	1	100,0	,0	,0	100,0
		2	,0	100,0	,0	100,0
		3	,0	2,0	98,0	100,0

a. 98,7% of original grouped cases correctly classified.

Figure 15 Standardized Canonical Discriminant Function Coefficients

	Function	
	1	2
Zscore(RDEXPPER)	,699	,538
Zscore(PATENTS)	,338	-,414
Zscore(GDPPC)	,449	,308
Zscore(POPUL)	,630	-,389

Figure 16 Canonical Discriminant Function

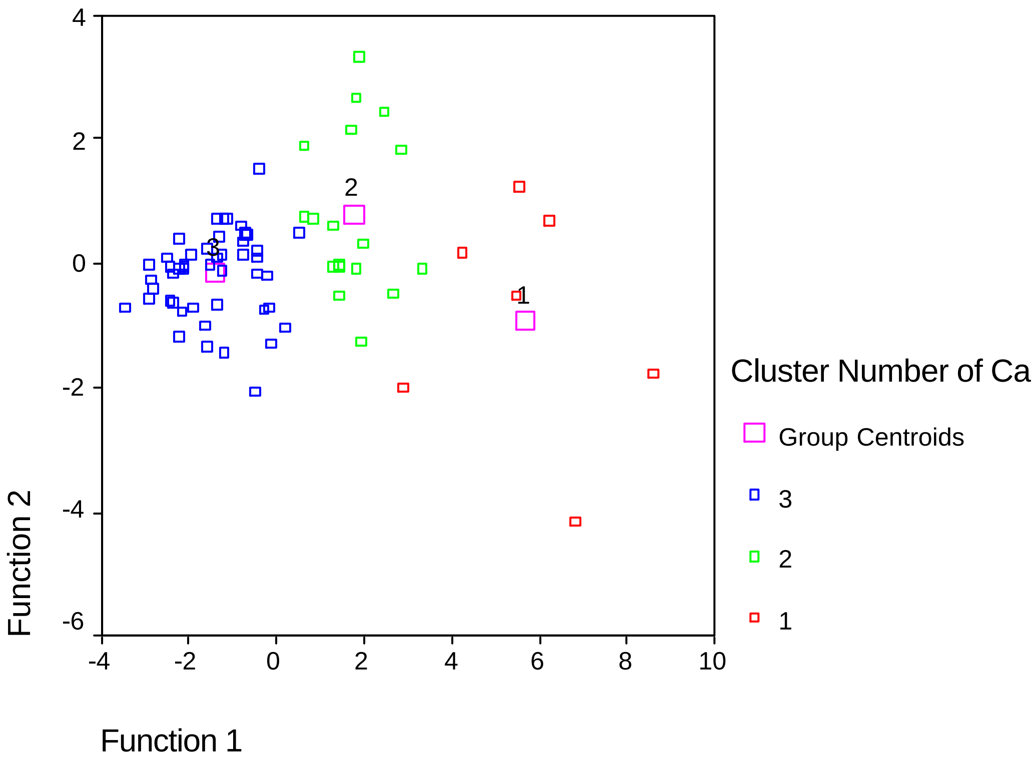


Figure 17 Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
Zscore(RDEXPPER)	,399	26,407	4	70	,000
Zscore(PATENTS)	,163	90,033	4	70	,000
Zscore(GDPPC)	,281	44,713	4	70	,000
Zscore(POPUL)	,438	22,431	4	70	,000

Figure 18 Wilk's Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 3	,038	229,215	12	,000
2 through 3	,303	83,486	6	,000
3	,838	12,398	2	,002

Figure 19 Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	7,019[a]	78,2	78,2	,936
2	1,761[a]	19,6	97,8	,799
3	,194[a]	2,2	100,0	,403

a. First 3 canonical discriminant functions were used in the analysis.

Figure 20 Variables in the Analysis

Step		Tolerance	F to Remove	Wilks' Lambda
1	Zscore(PATENTS)	1,000	90,033	
2	Zscore(PATENTS)	,995	68,482	,281
	Zscore(GDPPC)	,995	32,351	,163
3	Zscore(PATENTS)	,992	51,806	,153
	Zscore(GDPPC)	,990	25,510	,095
	Zscore(RDEXPPER)	,992	8,430	,057

Figure 21 Classification results (a)

			Predicted Group Membership					
		Cluster Number of Case	1	2	3	4	5	Total
Original	Count	1	13	0	0	0	2	15
		2	0	3	0	0	0	3
		3	0	0	2	0	0	2
		4	0	0	0	24	2	26
		5	1	0	0	0	28	29
	%	1	86,7	,0	,0	,0	13,3	100,0
		2	,0	100,0	,0	,0	,0	100,0
		3	,0	,0	100,0	,0	,0	100,0
		4	,0	,0	,0	92,3	7,7	100,0
		5	3,4	,0	,0	,0	96,6	100,0

a. 93,3% of original grouped cases correctly classified.

Figure 22 Misclassified Regions

Region	Code	Non-hierarchical Cluster	Discriminant
Champagne-Ardenne	FR21	4	5
Basse-Normandie	FR25	4	5
Berlin	DE3	1	5
Niedersachsen	DE9	1	5
Sachsen-Anhalt	DEE	5	1

Figure 23 Standardized Canonical Discriminant Function Coefficients

	Function		
	1	2	3
Zscore(RDEXPPER)	,414	,320	,857
Zscore(PATENTS)	,754	-,632	-,200
Zscore(GDPPC)	,452	,790	-,427

Figure 24 Canonical Discriminant Function

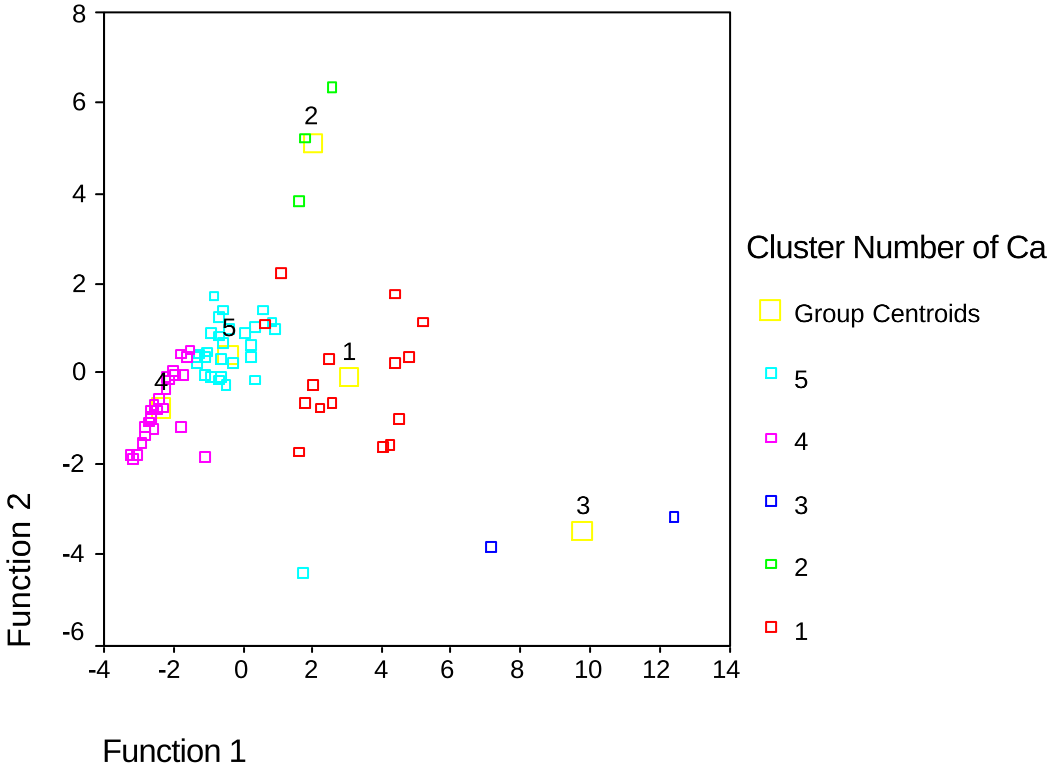


Figure 25 Data Set Used and Grouping Obtained (5 clusters)

Region	NUTS2 Code	Discriminant	RDExpPercentage	NumberPatents	Patents wrt average	% Patents	GDPpc	Population
Austria	AT	2,00	1,30	1.692,00	251,90	3,36	115,44	7.913.800,00
Arr. Admin. Bruxelles-Capitale - Admin. Arr. Brussel hoofdstad	BE1	2,00	1,70	395,00	58,81	0,78	168,50	941.200,00
Vlaams Gewest	BE2	2,00	1,60	1.463,00	217,81	2,90	107,16	5.771.200,00
Region Wallonne	BE3	3,00	1,40	545,00	81,14	1,08	87,70	3.269.100,00
Baden-Württemberg	DE1	1,00	3,60	2.057,00	306,24	4,08	142,15	10.075.300,00
Bayern	DE2	1,00	2,80	2.056,00	306,09	4,08	130,67	11.683.100,00
Berlin	DE3	2,00	3,40	301,00	44,81	0,60	109,50	3.455.900,00
Brandenburg	DE4	3,00	1,50	234,00	34,84	0,46	56,10	2.543.800,00
Bremen	DE5	2,00	2,80	24,00	3,57	0,05	173,90	684.500,00
Hamburg	DE6	2,00	1,90	198,00	29,48	0,39	215,70	1.678.800,00
Hessen	DE7	2,00	2,10	2.296,00	341,82	4,56	150,53	5.878.200,00
Mecklenburg-Vorpommern	DE8	3,00	1,10	20,00	2,98	0,04	52,40	1.867.700,00
Niedersachsen	DE9	2,00	1,70	630,00	93,79	1,25	116,93	7.526.700,00
Nordrhein-Westfalen	DEA	1,00	1,70	4.147,00	617,39	8,23	124,46	17.594.500,00
Rheinland-Pfalz	DEB	2,00	1,90	1.719,00	255,92	3,41	111,20	3.851.100,00
Saarland	DEC	3,00	0,90	76,00	11,31	0,15	121,50	1.080.400,00
Sachsen	DED	2,00	2,30	1.526,00	227,19	3,03	76,13	4.596.100,00
Sachsen-Anhalt	DEE	3,00	1,50	2.255,00	335,72	4,48	51,80	2.809.200,00
Schleswig-Holstein	DEF	3,00	1,10	208,00	30,97	0,41	116,40	2.664.100,00
Thüringen	DEG	3,00	1,80	586,00	87,24	1,16	49,90	2.552.100,00
Noroeste	ES1	3,00	0,50	22,00	3,28	0,04	66,07	2.792.800,00
País Vasco	ES21	3,00	1,20	24,00	3,57	0,05	86,90	2.102.300,00
Comunidad Foral de Navarra	ES22	3,00	0,80	8,00	1,19	0,02	91,80	521.700,00
La Rioja	ES23	3,00	0,40	1,00	0,15	0,00	85,40	260.300,00
Aragón	ES24	3,00	0,60	14,00	2,08	0,03	82,40	1.207.300,00
Comunidad de Madrid	ES3	2,00	1,80	164,00	24,42	0,33	93,90	4.910.200,00
Castilla y León	ES41	3,00	0,50	23,00	3,42	0,05	65,00	2.618.200,00
Castilla-la Mancha	ES42	3,00	0,40	4,00	0,60	0,01	61,80	1.716.700,00
Extremadura	ES43	3,00	0,30	1,00	0,15	0,00	49,60	1.131.200,00
Cataluña	ES51	3,00	0,90	352,00	52,40	0,70	91,90	6.018.200,00
Comunidad Valenciana	ES52	3,00	0,50	32,00	4,76	0,06	74,80	3.797.700,00
Andalucía	ES61	3,00	0,60	29,00	4,32	0,06	56,00	6.983.700,00
Región de Murcia	ES62	3,00	0,50	2,00	0,30	0,00	68,60	1.038.100,00
Canarias	ES7	3,00	0,40	3,00	0,45	0,01	72,10	1.501.700,00
Finland	FI	2,00	2,30	1.351,00	201,13	2,68	102,99	5.042.000,00
Île de France	FR1	1,00	3,30	5.751,00	856,19	11,42	172,30	10.785.000,00
Champagne-Ardenne	FR21	3,00	0,40	144,00	21,44	0,29	103,60	1.348.400,00
Picardie	FR22	3,00	1,00	410,00	61,04	0,81	92,00	1.794.000,00
Haute-Normandie	FR23	3,00	1,20	334,00	49,73	0,66	111,10	1.756.900,00
Centre	FR24	3,00	1,30	342,00	50,92	0,68	102,80	2.398.400,00
Basse-Normandie	FR25	3,00	0,70	99,00	14,74	0,20	95,60	1.400.300,00
Bourgogne	FR26	3,00	0,90	199,00	29,63	0,40	97,20	1.574.600,00
Nord - Pas-de-Calais	FR3	3,00	1,20	477,00	71,01	0,95	91,60	3.974.700,00
Lorraine	FR41	3,00	0,80	221,00	32,90	0,44	95,10	2.309.600,00
Alsace	FR42	3,00	1,10	645,00	96,03	1,28	114,70	1.655.500,00
Franche-Comté	FR43	3,00	2,00	52,00	7,74	0,10	97,30	1.103.900,00
Pays de la Loire	FR51	3,00	0,80	194,00	28,88	0,39	97,50	3.097.900,00
Bretagne	FR52	3,00	1,40	181,00	26,95	0,36	92,30	2.704.200,00
Poitou-Charentes	FR53	3,00	0,60	158,00	23,52	0,31	89,90	1.492.600,00
Aquitaine	FR61	3,00	1,30	531,00	79,05	1,05	99,30	2.830.900,00
Midi-Pyrénées	FR62	2,00	3,20	453,00	67,44	0,90	93,60	2.469.100,00
Limousin	FR63	3,00	0,60	39,00	5,81	0,08	87,40	720.300,00
Rhône-Alpes	FR71	2,00	2,20	2.428,00	361,47	4,82	110,10	5.454.200,00
Auvergne	FR72	3,00	1,40	173,00	25,76	0,34	88,90	1.316.700,00
Languedoc-Roussillon	FR81	3,00	1,50	416,00	61,93	0,83	85,30	2.164.500,00
Provence-Alpes-Côte d'Azur	FR82	2,00	2,00	428,00	63,72	0,85	101,50	4.341.900,00
Attiki	GR3	3,00	0,70	28,00	4,17	0,06	50,90	3.514.200,00
Rest of Greece	GRR	3,00	0,50	25,00	3,72	0,05	45,65	4.526.300,00
Nordovest	IT1	2,00	1,50	783,00	116,57	1,55	121,80	6.089.300,00
Lombardia	IT2	1,00	1,20	2.640,00	393,04	5,24	130,20	8.895.900,00
Nord Est	IT3	3,00	0,60	783,00	116,57	1,55	119,73	6.407.700,00
Emilia-Romagna	IT4	3,00	0,80	672,00	100,05	1,33	128,00	3.915.600,00
Centro	IT5	3,00	0,80	468,00	69,67	0,93	104,70	5.772.200,00
Lazio	IT6	2,00	1,90	518,00	77,12	1,03	117,00	5.151.900,00
Abruzzo-Molise	IT7	3,00	0,70	41,00	6,10	0,08	85,00	1.583.600,00
Campania	IT8	3,00	0,80	73,00	10,87	0,14	68,80	4.579.200,00
Sud	IT9	3,00	0,40	39,00	5,81	0,08	65,57	5.916.700,00
Sicilia	ITA	3,00	0,50	46,00	6,85	0,09	69,90	4.981.900,00
Sardegna	ITB	3,00	0,60	21,00	3,13	0,04	78,60	1.649.300,00
Noord Nederland	NL1	3,00	1,20	174,00	25,90	0,35	103,20	1.201.000,00
Oost-Nederland	NL2	2,00	2,20	726,00	108,09	1,44	86,17	3.108.200,00

6 References

Achilladelis, B., Schwarzkopf, A. (1990): "The Dynamics of Technological Innovation: The Case of the Chemical Industry", *Research Policy*, Vol. 19, pp.1-32.

Albach, H., Audretsch, D.B., Fleischer, M., Greb, R., Höfs, E., Lars-Hendrik, R., Schulz, I. (1996): "Innovation in the European Chemical Industry", *Wissenschaftszentrum Berlin für Sozialforxchung* (WZB). Internet: http://skylla.wz-berlin.de/pdf/1996/iv96-26.pdf

Audretsch, D.B., Feldman, M.P. (1996): "R&D Spillovers and the Geography of Innovation and Production", *The American Economic Review*, Vol. 86, Iss. 3, June, 630-640.

CEFIC (2001): "Facts and Figures. The European Chemical Industry in a Worldwide Perspective 2000".

European Commission (1995): "Green Paper on Innovation", Brussels. Internet: http://europa.eu.int/en/record/green/gp9512/ind_inn.htm

-(1996): "The First Action Plan for Innovation in Europe", Brussels. Internet: http://www.cordis.lu/innovation-smes/src/lib-iap.htm

-(2000a): "Towards a European Research Area. Science, Technology and Innovation. Key Figures 2000", Brussels. Internet: http://europa.eu.int/comm/research/pdf/keyfiguresihp.pdf

-(2000b): "European Trend Chart on Innovation", Brussels.

-(2001): "Second Cohesion Report", Brussels. Internet: http://europa.eu.int/comm/regional_policy/sources/docoffic/official/reports/contentpdf_en.htm

-(2002): "Commission Recommendation for the 2002 Broad Guidelines of the Economic Policies of Member States and the Community", Brussels. Internet: http://europa.eu.int/comm/economy_finance/publications/european_economy/2002/ee402en.pdf

Fleischer, M., Kelm, S., Palm, D. (2000): "Regulation and Innovation in the Chemical Industry", *Institute for Prospective Technological Studies*.

Gambardella, A., Malerba, F. (eds.) (1999): "The Organization of Economic Innovation in Europe", Cambridge University Press.

García-Fontes, W., Gambardella, A. (1996): "Regional Linkages through European Research Funding", *Economics of Innovation and New Technology*, 1996, v. 4, iss. 2.

García-Fontes, W., Geuna, A. (1999): "The Dynamics of Research Networks in Europe", in *The Organization of Economics Innovation in Europe*, Gambardella, A., Malerba. F. (eds.), Cambridge University Press, Chapter 14, pp. 343-366.

Griliches, Z. (1979): "Issues in assessing the contribution of research and development to productivity growth", *The Bell Journal of Economics*, Vol. 10 (1), pp. 92-116

-(1990): "Patent Statistics as Economic Indicators: A Survey", *Journal of Economic Literature*, Vol. XXVIII, December, pp. 1661-1707. Internet: http://links.jstor.org/sici?sici=0022-0515%28199012%2928%3A4%3C1661%3APSAEIA%3E2.0.CO%3B2-2

Jaffe, A., Trajtenberg, M. Henderson, R. (1993): "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *Quarterly Journal of Economics*, 63 (3), 577-598.

OECD (1996): "The Knowledge-Based Economy", *OECD*, Paris.

Rosenkranz, S. (1995): “Simultaneous Choice of Process and Product Innovation”, mimeo, *Wissenschaftszentrum Berlin für Sozialforxchung* (WZB).

Stoneman, P. (ed.) (1995): “Handbook of the economics of innovation and technological change”, Blackwell