

Regional Linkages through European Research Funding ¹

Alfonso Gambardella
Istituto di Studi Aziendali, Università di Urbino and
IEFE, Università Bocconi, Milano

and

Walter Garcia-Fontes
Universitat Pompeu Fabra

October 1994

¹We gratefully acknowledge the useful comments and suggestions of Aldo Geuna, Michael Greenacre, Paul David, Albert Satorra, Luc Soete, and Ed Steinmueller, as well as of seminar participants at the Maastricht Institute for Research in Innovation and Technology (MERIT), Universitat Pompeu Fabra and the 1993 Meetings of the European Economic Association at Maastricht, the Netherlands. We thank MERIT's hospitality while we were conducting this research, as well as the financial support of the European Union Human Capital and Mobility Program, Contracts Nr. ERBCHRXT920002 and Nr. ERB4050PL930320. Alfonso Gambardella acknowledges the financial support from the Italian Ministry of University and Research.

Abstract

We use cluster analysis to describe the regional linkages that arise through funding of research contract networks in the EU. We find five significantly different kinds of networks, that we label as follows: 1) Technological development, 2) Basic research, 3) Quasi-elite, 4) Elite and 5) Southern. These networks are described in terms of three basic dimensions: quality, type of partners, and size combined with cost of the project. In terms of variables constructed along these dimensions, we find that the networks are homogeneously formed and that regions of similar technological capabilities are linked together. We discuss this empirical fact by means of a model in which researchers are matched by skill.

JEL Classifications: O32 (Management of Technological Innovation and R & D), O38 (Government Policy).

1 Introduction

The allocation of public funds to scientific and technological research presents an inherent trade-off. If public research agencies seek to maximize the output (publications, patents) of the particular program they are managing, they will allocate more resources to applicants with greater expected productivity. In science, expected productivity is highly correlated with past performance (e.g. Lotka, 1926; Allison et al., 1976). This may give rise to phenomena of increasing returns in the pattern of allocations over time. Scientists or groups with better past records obtain more funds in subsequent periods. The level of funding influences their performance, and hence their ability to obtain more funds in the future – the “Matthew effect” (Merton, 1968; see also Dasgupta and David, 1992, and David, 1993).

If agencies take a longer term perspective, they would have to take into account the effects of their actions on future research programs as well. Static efficiency may in fact magnify differences in performance – among groups, people, or regions – more than what is implied by initial conditions, “innate” abilities, or research potential. Agencies would then want to allocate part of the funds to applicants with lower expected productivity who could nonetheless grow significantly in the future. This, however, implies that they sacrifice some short-term output.

The allocation of research funds by the European Union (EU) is one of the most apparent examples of this trade-off. On the one hand, the EU wants to maximize present research performance. This is important for enhancing European science and technology, and for industrial competitiveness. In the light of intense international competition, it may be too costly, even from a social point of view, to sacrifice short-term research outcomes in favor of longer term goals. On the other hand, the objectives of European integration highlight the importance of spreading research opportunities across many regions.

Relatedly, as many authors have pointed out (particularly, Rosenberg, 1982), economic rents do not really arise from the generation of new innovations and technologies, but from their development and commercialization on a relatively large scale. And successful (economic) development and commercialization of innovations often depend on the interaction with potential

users and larger markets. Interactions with less developed regions of the European Union is then an important means of exploiting a truly continental market for development and commercialization of innovations.

On the other hand, there have been some recent efforts to look at the effects of geographical proximities on research linkages. Jaffe (1989) shows that the elasticity of academic research expenditures on corporate patents is higher the more the patent comes from a firm (or institution) that is geographically close to where the academic money is spent. Moreover, Henderson et. al. (1993) find that citations in patents tend to be geographically localized, i.e. citations to domestic patents are more likely to be domestic and more likely to come from the same state as the cited patent.

This paper focuses on the contracts signed in 1990 in the EU Brite-Euram (henceforth BE) program. BE is an interesting program for our purposes. First, it covers many technologies (new materials, chemicals and chemical processing, aeronautics, industrial automation, simulation, etc.), and it comprises a heterogeneous set of participants (large firms, small-medium firms, universities, other research centers). It is then representative of a wide set of institutional types, technologies, regions, and industries. Moreover, BE is concerned with both the generation of new technologies and their development and commercialization.

Each BE contract is assigned to a network of participants. The purpose of this paper is to describe the composition of these networks using cluster analysis, and provide an interpretation of the clusters of networks and relevant variables using correspondence analysis (Do we observe mixing of firms with universities ? High research-intensive regions with low-intensive ones ?). These multivariate techniques, cluster and correspondence analysis, are purely descriptive tools. However, they may help in organizing the data in ways that may point to structural models that have predictive and testable elements.

The paper is organized as follows. The next section describes BE, the variables in our data set, and the clusters. Section 3 attempts some interpretation of our results. Section 4 presents the conclusions.

2 The 1990 Brite–Euram Contracts

2.1 The Brite-Euram Program

The 1990 BE contracts regard one of the four years of this program (1989-1992). The main objective of BE 1989-1992 was to enhance the competitive position of the Community’s manufacturing industries. Related goals included trans-frontier collaboration in strategic industrial research and the transfer of technology across Community frontiers and between sectors, particularly those with many small-medium enterprises (SME) (CEC, 1993). Although not an explicit objective, European “cohesion”, i.e. stronger inter-relationships among most and least favored regions of the Community, was mentioned to be a desirable outcome of the program (CEC, 1993, p.12). This suggests that allocation of resources in this program did take into account longer term objectives as well.

We obtained our data from DGXII (1991), which lists all contracts signed by DGXII in 1990. For each contract (network), DGXII (1991) provides the following information: contract number, title of project, name and location (ISO regions) of partners, type of institution (large firm, small-medium firm, university, research center, other), its position in the network (main contractor, secondary contractor, third contractor, sub-contractor of main contractor, of secondary contractor, etc.), duration of project (in years), total cost of the project, total EC contribution, break-down of costs and EC contribution for each participant in the networks ¹.

We matched these data with the number of European patents (from the European Patent Office) by regions in three main technological classes for the years 1978–1990. We tried to approximate the technological capabilities of a particular region by three different kind of measures.

In the first place, we were interested in an output measure, so we did a patent count by region. The address of the main inventor provided us the region of origin of the patent, in the case of multiplant corporations that operate in different European regions.

In the second place we were interested in a specialization measure, adapted

¹Total EC contribution amounts to 100 % of cost for non-profit institutions, and 50 % for firms.

to the character of the BE program. Three classes were created after aggregating homogeneous EPO technological sub-classes. Our three classes – new materials, aircraft, mechanical engineering – roughly correspond to the technologies targeted by the BE program.

Finally we wanted to have a capabilities measure for each region. Consequently we build our patent count starting from several years prior to the years of the Brite–Euram program and continuing through the period.

Our patent count by regions represents an approximate measure of the regional technological capabilities, in the main technological classes of BE. The use of patent data has some drawbacks, for instance it is known that an important part of the output of research cannot be patented, especially for basic research. In any case here we are not trying to measure regional technological capabilities *per se*, but only to discriminate among European regions and see how these regions are linked by means of the BE program.

For the large European countries, Italy, France, UK, Germany, and Spain our regions correspond to the political regions or provinces (e.g., for Italy, Lombardy, Piedmont, etc.; for France, Alsace, Aquitaine, etc.; for the UK, East Anglia, Yorkshire, etc.; for Germany, Baden- Wuerttemberg, Berlin, etc.; for Spain, region of Madrid, of Barcelona, etc.). For the small countries, Belgium, the Netherlands, Luxembourg, Denmark, we use the patents of the country as a whole. This is justified by the fact that the size of these countries (e.g. in terms of population) is roughly equivalent to that of the main regions of the largest countries, such as the Paris region in France, Bayern or Nordrhein–Westfahlen in Germany or the North-West region in the United Kingdom. We tried some disaggregation by region on the small countries, but this created problems of comparability with the regions of the larger countries. For Greece we distinguished between the Athens region and the rest of the country, while for Portugal we did not distinguish by regions. Although geographically large, these countries have a very small number of patents, and further regional distinctions produce too fine a partition. As we shall discuss below, we used our regional patent data to create ten classes of regional technological capabilities, rather than using them as absolute numbers. This further reduces the problems of regional comparability.

In the Appendix, Table 9 presents the main figures describing the regions in Europe that contribute participants to the BRITE/EURAM program.

The first column of the table presents the number of patents in the three technological classes chosen, produced by firms located in each region. The second column shows the number of participants coming from each region that act as main contractor. Comparing these two columns, it can be seen that regions with high number of patents usually contribute a higher number of main contractors to the networks. This is the case for the Paris region with 2,104 patents and 16 main contractors, Holland with 525 and 13, or South East region of the United Kingdom with 703 and 13. The pattern is not so clear for the third column, where the total number of participants contributed by the region is reported. In this case the heterogeneity is much higher. Although the Paris region is again the region with the highest number of participants, 90, we can find cases as the Madrid region in Spain with only 11 patents and 23 participants, or the Bayern region in Germany with 1,046 patents and 29 participants. The table is completed with information about the population of the regions and GDP per person.

2.2 Multivariate descriptive analysis of the networks

Our analysis is based on three different dimensions: a) the technological capability or quality of the network, based on the technological capabilities of the regions contributing participants to the networks and the region of origin of the main contractor, b) type of partners (private firms, or public educational institutions or research laboratories), based on the type of the main contractor and the composition of the network in terms of the proportion of private institutions in them, and c) the size and cost of the network.

In order to evaluate the technological capabilities of the regions contributing participants to each BE network, we gathered the regions in ten different groups of approximately equal size, where group 1 represents the regions with smallest number of patents and group 10 represents the regions with highest number of patents. The details for this grouping are shown in the Appendix. We use this grouping as a measure of the technological capability of the region and we call this measure *quality*, as it represents a discrete scalar indicator of the geographical technological capabilities for the region of origin of each participant. In our analysis we use the group of the main contractor as an indicator of her geographical characteristics, and an average formed over all the groups present in the network as an indicator of the

Table 1: Descriptive Statistics for the Quality, Size, Type and Cost of the Networks

Variable	Mean	Standard Deviation	Min.	25 % Quartile	Median	75 % Quartile	Max.	Range 75-25 %
Size of the network	5.17	2.26	1	4	5	6	14	2
Average Quality	8.03	1.43	3.8	7.21	8.14	9.2	10	2.00
Quality of the main contractor	8.61	2.23	1	8	10	10	10	2
Dummy for firms as main contractors	0.62	0.49	0	0	1	1	1	1
Proportion of firms within the network	0.54	0.33	0	0.34	0.6	0.8	1	1
Cost of the project (1,000 Ecus)	559.91	331.63	0	337.95	497.25	728.25	2,169	390.3
Fraction financed by the Commission	0.60	0.17	0.26	0.5	0.54	0.62	1	0.12

average geographical characteristics of the network.

We will base our analysis of the composition of the networks in the following variables: 1) Size of the network, 2) geographical quality of the main contractor, 3) average geographical quality of the network, 4) a dummy variable which takes the value 1 if the main contractor is a firm and 0 otherwise, 5) an index of privateness of the network, which is the proportion of firms over the total number of participants, 6) the total cost of the network and 7) the fraction financed by the European Commission through the BE program.

Table 1 provides descriptive statistics for the networks.

2.3 Clustering the networks

We use cluster analysis to analyze possible groupings according to the composition of the networks. Since our data contain continuous, categorical and binary variables, we cannot perform cluster analysis directly on them. The procedure chosen consists of first extracting principal components from the original variable, then performing cluster analysis on the principal components. Finally we use canonical discriminant analysis to represent the clusters on a plane, and correspondence analysis to visualize the association of our variables within and between the clusters suggested. The details are explained in the Appendix.

We chose five clusters to represent our data, since this is the number of clusters that minimizes our clustering criterion. The clusters are represented graphically using canonical discriminant analysis in Figure 1 in the Appendix.² The figure shows that there is little overlapping between the individuals of the five clusters. We will label the clusters as follows:

Cluster 1: Technological Development

Cluster 2: Basic Research

Cluster 3: Quasi-Elite

Cluster 4: Elite

Cluster 5: Southern

Below, in the discussion about the characteristics of the clusters of networks, we will justify this labeling.

The difference in composition of the clusters in terms of each variable can be visualized by using correspondence analysis on the frequency tables for each variable. Correspondence analysis is a multivariate statistical technique used to study the association of variables within a contingency table. We will study the contingency tables obtained and describe the main characteristics of the clusters uncovered in the previous step.

²Canonical discriminant analysis is used here to reduce the dimensionality of the data set and plot the clusters on a two dimensional space.

Table 2: Regional quality of the main contractor

Cluster	Quality group of the main contractor										Sum
	1	2	3	4	5	6	7	8	9	10	
Basic Research	1	1	0	0	1	2	3	4	5	12	29
Elite	0	0	0	0	0	0	0	0	6	21	27
Quasi Elite	0	0	0	0	0	2	1	3	4	13	23
Southern	1	3	3	1	1	0	0	0	0	0	9
Technological Development	0	0	0	1	2	4	2	3	11	32	55
Sum	2	4	3	2	4	8	6	10	26	78	143

We begin by discussing the geographical quality of the network, according to the definition provided in the previous section. Table 2 shows how the clusters are composed in terms of the regional group of the main contractor, or director of the project. In Figure 2 in the Appendix we present the symmetric correspondence map for this table. Inspecting this map, we can see that there is association between the Technological Development and Elite clusters and main contractors of quality groups 9 and 10, between the Basic Research and Quasi Elite clusters and main contractors 6, 7 and 8, noting that the Quasi Elite cluster also shows association with the high quality group, and between the Southern cluster and the low quality main contractors. In other words, in terms of the quality of the main contractor, the Technological Development and Elite clusters are characterized by main contractors coming from technologically advanced regions. The Southern cluster is composed by networks co-ordinated by main contractors coming from less “patent-intensive” regions. In Table 3 and Figure 3 we present the composition of the clusters in terms of average quality of the partners involved. The Technological Development and Elite cluster are associated with high average quality (in terms of number of regional patents); that is, partners also come from technologically advanced regions. In the Basic Research and Quasi Elite clusters, the average quality of the partners’ regions is spread more evenly.

Table 3: Average regional quality of the network

Cluster	Average quality of the networks						Sum
	0-5	6	7	8	9	10	
Basic Research	2	3	6	7	5	6	29
Elite	0	0	1	4	7	15	27
Quasi Elite	0	4	4	4	9	2	23
Southern	3	3	3	0	0	0	9
Technological Development	0	4	3	22	13	13	55
Sum	5	14	17	37	34	36	143

These clusters include networks with partners coming from high, medium or low patent- intensive regions. The Southern cluster is composed of partners coming from low patent-intensive regions. As we shall discuss in more detail, main contractors coming from less technologically advanced regions seem to be unable to attract partners coming from more advanced regions.

In Table 4 and Figure 4 we analyze the association between the clusters and the type of the main contractor (big or small-medium firm, universities, etc.). In terms of this variable, the Technological Development cluster is associated with networks whose main contractors are either big or small-medium firms. In the Basic Research cluster, main contractors are non-profit institutions (university or research centers). The Quasi Elite and Elite clusters, show greater spread, even though large firms are main contractors in most of the cases. In the Southern cluster, main contractors are either firms (small or large) or government research centers, but not universities. In less advanced regions, universities do not take the initiative.

As far as privateness is concerned, Table 5 and Figure 5 show that the Technological Development cluster is very “private”, whilst the Basic Research cluster is very “public”. The Elite cluster shows an association with the 0.2-0.4 level, which indicates that there is an important proportion of

Table 4: Type of the Main Contractor

Cluster	Proportion of firms							Sum
	big	edu	oth.	rmx	rpr	rpu	sme	
Basic Research	0	21	0	1	0	7	0	29
Elite	9	2	2	2	3	4	5	27
Quasi Elite	12	4	1	1	2	0	3	23
Southern	3	0	0	4	0	0	2	9
Technological Development	36	0	1	0	0	0	18	55
Sum	60	27	4	8	5	11	28	143

Table 5: Proportion of firms within the networks

Cluster	Proportion of firms					Sum
	0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1	
Basic Research	29	0	0	0	0	29
Elite	0	9	7	7	4	27
Quasi Elite	2	3	6	10	2	23
Southern	0	1	5	1	2	9
Technological Development	0	3	12	16	24	55
Sum	31	16	30	34	32	143

Table 6: Size of the networks

Cluster	Number of participants						Sum
	0-2	3-4	5-6	7-8	9-10	> 10	
Basic Research	7	10	11	1	0	0	29
Elite	3	8	15	1	0	0	27
Quasi Elite	0	0	6	8	5	4	23
Southern	0	0	3	3	3	0	9
Technological Development	4	24	24	3	0	0	55
Sum	14	42	59	16	8	4	143

public partners in these networks. The Quasi-Elite and Southern clusters show a fairly even distribution among private and public networks.

Tables 6 and 7 and Figures 6 and 7 present the analysis in terms of the size of the network and the total cost involved. In terms of these variables, the Technological Development, Basic Research and Elite clusters are associated with networks with smaller number of participants, whereas the Quasi-Elite and Southern clusters are larger. But while the Technological Development and Basic Research clusters undertake relatively low cost projects, the Elite cluster is composed of networks performing large projects (in terms of costs). The Quasi-Elite cluster (high size) is low cost, and the Southern cluster is average cost.

3 Characteristics of the clusters

In this section we provide some initial interpretation of our results. Table 8 summarizes the characteristics of our five clusters.

As one can see from the Table, cluster 1 is the TECHNOLOGICAL DEVELOPMENT (TD) cluster. It is composed of main contractors coming

Table 7: Cost of the projects

Cluster	Total cost of the project (Thousand ECUs)					Sum
	0-250	250-500	500-750	750-1000	> 1000	
Basic Research	13	14	2	0	0	29
Elite	0	5	5	9	8	27
Quasi Elite	3	14	4	2	0	23
Southern	0	3	4	2	0	9
Technological Development	3	19	19	10	4	55
Sum	19	55	34	23	12	143

Table 8: Summary characteristics of the clusters

Clusters	Quality		Type		Size and Cost	
	Quality of main contractor	Average quality of network	Type of main contractor	Private- ness	Size	Cost
1. TECHNOL. DEVELOPMENT (38% of sample)	H	H(-)	B/S	PR	SM	SM/ SPREAD
2. BASIC RESEARCH (20%)	H	H(-)	E/R(-)	PB	SM	SM
3, QUASI- ELITE (16%)	H	H(-)	B/S(-) (spread)	PR(-)	BIG	SM
4, ELITE (19%)	H	H	SPREAD	SPREAD	SM	HIGH
5. SOUTHERN (6%)	L	L	B/S/R AV(PR)	AV/BIG	AV	

H:high,L:low,AV:average,SM:small

B:big firm, S:small-medium firm, E:educational, R:public lab

from high-tech regions, and of partners coming in good part from high-tech regions. It is co-ordinated by firms (either large or small), and it is very private. It is also small in size and in cost. The networks belonging to this cluster appear to have gathered around a well defined project of technological development. The privateness of the networks, and the fact that it is co-ordinated by firms, suggest that these are mainly projects aimed at developing particular technologies which may have some medium to short term economic applicability.

Cluster 2 is the BASIC RESEARCH (BR) cluster. Its networks gather basically non-profit institutions (universities or government research centers). Main contractor and partners come from high-tech regions. Clusters 1 and 2 then distinguish between downstream and upstream research networks. Moreover, there is little mixing of private and public institutions in the two, suggesting that profit and non-profit institutions, at least those financed by EC research programs, do seem to separate themselves in terms of the degree of “basicness” of their research. (This is reinforced by the consideration that these two clusters account for more than half of the total number of networks in our sample.) Less technologically advanced regions are excluded by both types of networks.

Cluster 4 has also interesting features. Main contractor and partners come from high quality regions. The type of the main contractor and the degree of privateness is evenly distributed. In other words, main contractors of different types are equally likely, and the networks cannot be identified precisely to be fully private or public (networks are evenly distributed in terms of privateness). Finally, these networks include a small group of participants, and they undertake costly projects. We named cluster 4 the ELITE (E) cluster. The composition of these networks suggest that these are networks gathering “sophisticated” institutions, and participants appear to be linked more by their belonging to a common, “elite”, scientific or technological “club”, than by anything else. The even distribution of types (both of main contractor and of participants) suggest that these “elite” groups can be found in different types of institutions (whether non-profit or firms). Moreover, the good mixing of universities and firms suggest that these networks undertake projects that are somewhat more basic than those of the Technological Development cluster, and more applied than those of the Basic Research cluster. A related consideration is that, unlike the first two clusters (and cluster 5, as we shall

see below), these are rather heterogeneous networks, in the sense that they are formed by heterogeneous groups (apart from the fact that all groups come from advanced regions).

Cluster 3 is similar to cluster 4, but we suggest that it is probably less “elitist” (the QUASI-ELITE cluster – QE). Main contractors coming from advanced regions match with partners coming from other advanced regions, even though partners in the ELITE cluster are distributed more exactly towards the upper end of the average quality distribution of technological classes. Type of main contractor is evenly distributed, but main contractors are more frequently firms (large or small) than in the Elite cluster. At the same time, other partners are more “private” than in the Elite cluster. These networks are also larger (in number of participants), and they are less costly (smaller projects). In sum, this cluster seems to have characteristics that are similar to the Elite cluster, but in the latter networks are more clearly identified (in terms of quality), and more evenly distributed (in terms of types). Also, the fact that projects are smaller, and participants are in greater number suggests that here participants are not really members of well established “clubs”. The networks in the Quasi-Elite cluster are probably networks that are attempting to become “elite”, but that have not yet completely reached this objective.

Cluster 5 has quite obvious features. These are networks whose main contractor and partners come from less advanced regions. We called it the SOUTHERN (S) cluster because participants presumably come mostly from southern regions of the Community. Another feature of these networks is that the initiative (main contractor) is taken by firms or government research centers (not universities). Universities are brought in as partners. Size and cost are evenly distributed.

4 Interpretation of the results

It is possible to interpret the characteristics of our clusters by looking at them in terms of their degree of homogeneity along our three dimensions: quality, type, and size/costs. Generally speaking, an important EU policy objective is to encourage heterogeneity in the formation of networks. This is especially true for the type of participants. Particularly, the EU aims at favoring

mixing of profit and non-profit research institutions to encourage transfer of knowledge between more upstream and more downstream research. Second, as the stated BE goals pointed out, another important objective was to favor transfer of technologies across geographical borders. This implies, among other things, encouraging interactions among the most and the least advanced regions. A final issue regards diffusion (versus generation) of technologies, or, to put it differently, the possibility of involving participants from larger geographical areas of the Community to encourage development and commercialization of innovations in a truly continental area. In essence, our results suggest that:

- (a) The BE 1990 program was rather successful in producing networks capable especially of generating new technologies. As suggested earlier, the TD and BR clusters cover almost 60 % of all networks in the sample. The fact that the BR networks are involved primarily in the generation of new knowledge is fairly obvious. Yet, we think that the same can be said about the TD networks, and this is especially because they do not really involve the large continental market of Europe, but only a well-defined set of advanced regions (in spite of the fact that they are primarily concerned with development, and hence possibly commercialization of new technologies). Less advanced regions are excluded not only from the upstream research networks, but also from the downstream ones.
- (b) This also suggests that the program was probably not very successful in encouraging diffusion of new technologies, and more generally diffusion of new technological capabilities, especially across North-South borders of the Community. Again, this can be seen by the fact that networks tend to be rather homogeneous in terms of patent-intensity of regions, with clusters 1-4 being composed by main contractors and partners coming from advanced regions, and cluster 5 by institutions from less advanced regions.³

³Cohen and Levintahl (1989) have suggested that the diffusion of technological capabilities comes from the absorption of capabilities by apt R & D partners, and a posterior rediffusion to partners with more reduced technological capabilities. This could help in explaining the matching of high quality partners among themselves, but not the matching of low quality partners.

- (c) The program also seemed to be successful in involving small-medium firms, which are often cut out from the R&D circuits of larger firms. Particularly, a good number of networks in the TD cluster, and to some extent in the BR and E clusters, have small-medium firms as main contractors. Yet, consistently with the discussion above, these smaller firms appear to come in any case from more advanced regions. This is particularly troublesome if one considers that less advanced regions tend to have relatively higher proportions of smaller firms (e.g. South of Italy).

As we discussed earlier, these results have two implications. First, the observed policy effects have been more akin to maximization of short-term output of the program (present performance) rather than directed towards longer run objectives. This may be an important factor in creating Matthew-effect phenomena across North-South regions of the Community. More generally, not only do these effects seem to create greater North-South divergence in the generation of new technologies (BR cluster), but also in their downstream development and commercialization (QE and E clusters, and especially cluster TD). The second important implication is that the creation and the development of these technologies do not seem to exploit the potential of involving, at relatively earlier R&D stages, a truly continental market. Clearly, we are not offering value judgements here. We only point out the trade-off. The program appeared to lean more towards present performance and it was less keen on longer-term goals.

Finally, the matching of main contractors coming from advanced regions with partners coming from advanced regions, and of South-South networks, is intriguing from a theoretical point of view. Kremer and Maskin (1994) discuss why, in present industrial development, we observe increasing segregation by skill among workers. They develop a model in which agents can be of two types, managers or workers, and each type can be of high or low quality. A firm is composed of one manager and one worker. They argue that good managers will tend to match with good workers, and low quality managers with low quality workers. The intuition is that a high quality worker has a higher opportunity cost of being a worker than a low quality worker. This is because he has greater potential of becoming a manager. A high quality manager increases the productivity of the worker (of either types), and she increases the productivity of high quality workers more than

that of low quality ones. Thus, a high quality worker when matched with a lower quality manager will have greater incentives to become a manager because of the greater opportunity cost involved.

A very similar effect is at work here. High quality institutions have greater opportunities of becoming main contractors of networks that have good possibilities of being financed by the Community, than institutions coming from less advanced regions. If they join main contractors from less advanced regions, their expected productivity in the project will be lower, and thus they will have fewer opportunities of being financed by the EC. Therefore, they would tend to join co-ordinators from more advanced regions, which would increase their productivity, with implied benefits. Put differently, if they join Southern main contractors, they face greater opportunity cost of not being main contractors themselves, and they will then choose to be main contractors. As a result, southern institutions encounter serious difficulties in attracting partners from northern regions.

5 Conclusions

This paper studied the composition of the networks financed by the EU Brite-Euram Program. Our data include all BE contracts signed by DG-XII in 1990 (the second year of this 1989-1992 program). We matched these data with the number of European patents (EPO) by EU regions (e.g. Lombardy, Alsace, East Anglia, Baden- Wuerttemberg, etc.) in the three main technological areas of BE (new materials, aircraft, mechanical engineering). We then constructed ten classes ranking EU regions from class 1 (low number of patents) to class 10 (highest number of patents). Each class represents one decile of the distribution of patents by region, and it is then composed of approximately equal number of regions.

We used cluster analysis to identify groups of networks with similar characteristics. We then used correspondence analysis to identify the main characteristics of the five clusters suggested by the previous step. Cluster 1 is a “technological development” cluster. It comprises networks that are mainly composed by firms (small or large). Cluster 2 is composed of networks of non-profit institutions, and these networks are most likely to conduct fairly upstream research. Clusters 3 and 4 show a good mixing of profit and non-

profit institutions, and they are most likely to perform projects that integrate upstream research with development and commercialization objectives. Clusters 1 to 4 are composed primarily by partners coming from advanced regions. Cluster 5, which amounts to only 6 % of all networks in our sample, is composed of partners coming from less advanced regions.

This program appeared to be especially effective in encouraging the formation of networks with good capabilities of generating innovations and new technologies. It was probably less effective in encouraging diffusion of these technologies (and of related knowledge bases and capabilities) especially across the North-South border, as suggested by the fact that there is little joint participations of partners from North and South regions in the same networks. Put differently, the program has been more sensitive to production of research outcomes in the relatively short-term vis-a- vis longer term goals of more evenly distributed technological capabilities among European regions.

References

- ALLISON, P. D., J. S. LONG, AND T. K. KRAUZE (1982), "Cumulative advantage and inequality in science," *American Sociological Review*, 47:615–625.
- CEE-DGXII (1991), "List of contracts signed in 1990," DGXII, Brussels.
- COHEN, W. M. AND D. A. LEVINTAHL (1989), "Innovation and learning: The two faces of R & D," *Economic Journal*, 99:569–96.
- COMMISSION OF THE EUROPEAN COMMUNITIES (CEC) (1993), "Evaluation of the Brite-Euram Programme (1989–1992)," DG XII, Research Evaluation Report Nr. 53, EUR 15070 EN, Brussels.
- DASGUPTA, P. AND P. A. DAVID (1993). "Toward a new economics of science," Working Paper 320, CEPR, Stanford University. Forthcoming in *Research Policy*.
- DAVID, P. A. (1993). "Positive feedbacks and research productivity in science: Reopening another black box," Working Paper 93-017, MERIT/Faculty of Economics.
- GREENACRE, M. (1994). "Correspondence analysis and its interpretation," in Blasius, J. and M. Greenacre, editors, *Correspondence Analysis in the Social Sciences*. Academic Press.
- HENDERSON, R., A. JAFFE, AND M. TRAJTENBERG (1993), "Geographical localization of knowledge spillovers as evidenced by patent citations," *Quarterly Journal of Economics*, 108:577–98.
- JAFFE, A. (1989), "Real effects of academic research," *American Economic Review*, 79:957–70.
- KREMER, M. AND E. MASKIN (1994), "Segregation by skill and rise in inequality," Paper presented at the NBER Summer Workshop, July.
- LOTKA, A. J. (1926), "The frequency distribution of scientific productivity," *Journal of the Washington Academy of Sciences*, 16(12):317–323.

MERTON, R. K. (1968), "The Matthew effect in science," *Science*, 159:56–63.

ROSENBERG, N. (1982). *Inside the black box: technology and economics*. Cambridge University Press, Cambridge.

WARD, J. H. (1963), "Hierarchical grouping to optimize and objective function," *Journal of the American Statistical Association*, 58:236–244.

Appendix

A Regions and their contribution to the Brite-Euram networks

In Table 9 we present the main characteristics of the regions considered in this study. The patents reported correspond to the patents authored by firms located in the region. The participants of networks can be educational institutions, public research laboratories or firms, big or small-medium. We complete the table with information about the population and GDP per person.

B Technological capabilities groups

The grouping of regions by patent-intensity of the firms located in the region was based on the deciles of the distribution of all the regions. The period considered was 1978–1990 and the three classes chosen were new materials, aircraft and mechanical engineering. The number of patents for each group are the following:

- Group 1 (0 patents)
- Group 2 (more or equal to 1 and less than 4 patents)
- Group 3 (more or equal to 4 and less than 8 patents)
- Group 4 (more or equal to 8 and less than 16 patents)
- Group 5 (more or equal to 16 and less than 28 patents)
- Group 6 (more or equal to 28 and less than 40 patents)
- Group 7 (more or equal to 40 and less than 58 patents)
- Group 8 (more or equal to 58 and less than 113 patents)
- Group 9 (more or equal to 113 and less than 277 patents)
- Group 10 (more or equal to 277)

Table 9: Characteristics of the Regions in Europe with participants in the BRITE/EURAM program 1990

Code	Region	Number of Patents	Number of Main Contractors	Number of Participants in Networks	Population (thousands)	GDP per person ECU _s - 1985 (EUR12=100)
	Belgium	155	5	41	9,967.4	103
	Denmark	133	6	26	5,139.9	145
	France					
fra	Alsace	18	0	7	1,627.6	122
frb	Aquitaine	10	0	2	2,803.0	116
frc	Auvergne	0	1	2	1,321.4	94
frj	Languedoc Roussillon	1	0	2	2,124.8	93
frk	Limousin	2	1	1	722.6	93
frl	Lorraine	43	3	6	2,305.4	102
frn	Nord Pas De Calais	22	0	3	3,966.8	99
frp	Haute- Normandie	15	0	1	1,740.8	129
frq	Paris (Region)	2104	16	90	10,692.0	180
frr	Pays de Loire	27	1	6	3,064.6	102
frs	Picardie	22	0	3	1,814.2	105
fru	Provence, Cote d'Azur, Alpes	15	0	1	4,273.6	110
frv	Rhone Alpes	160	5	30	5,368.0	120
	Germany					
de01	Baden- Wuerttemberg	1,018	5	33	9,726.2	136
de02	Bayern	1,046	7	29	11,334.8	129
de03	Bremen	33	2	6	679.2	166
de04	Hamburg	70	1	4	1,640.9	212
de05	Hessen	503	6	16	5,717.0	146
de06	Nidersachsen	222	6	15	7,340.4	111
de07	Nordrhein- Westfalen	1,472	6	32	17,243.6	126
de08	Rheiland- Pfalz	209	0	2	3,733.8	114
de09	Saarland	31	0	1	1,070.2	121
de10	Schleswig- Holstein	65	0	1	2,614.2	108
de11	Berlin	32	1	5	3,420.2	146

Table 9: Characteristics of the Regions in Europe with participants in the BRITE/EURAM program 1990

Code	Region	Number of Patents	Number of Main Contractors	Number of Participants in Networks	Population (thousands)	GDP per person ECUS - 1985 (EUR12=100)
Greece						
gr35	Athens	5	1	14	3,506.4	36
	Rest of the country	3	0	8	6,582.3	42 ^a
Holland						
		525	13	43	14,951.5	111
Italy						
it04	Campania	5	1	4	5,831.4	66
it05	Emilia-Romagna	52	1	4	3,952.2	120
it06	Friuli-Venezia Giulia	16	0	2	1,202.0	107
it07	Lazio	13	0	3	5,181.0	101
it08	Liguria	7	1	3	1,723.2	122
it09	Lombardia	203	2	12	8,925.8	122
it12	Piemonte	133	1	9	4,356.8	115
it13	Puglia	6	0	3	4,075.4	68
it16	Toscana	8	0	3	3,561.6	106
it20	Veneto	27	0	2	4,391.6	97
	Ireland	32	3	15	3,502.8	67
	Luxemburg	32	2	2	391.8	121
Portugal						
		0	1	27	9,868.4	27
Spain						
	Andalucía	0	0	1	6,919.8	42
	Cantabria	0	0	1	527.2	59
	Cataluña	24	3	8	6,007.6	61
	C. Valenciana	0	0	3	3,786.6	54
	Madrid	11	2	23	4,877.8	62
	Navarra	2	0	1	521.2	71
	País Vasco	11	3	15	2,129.2	66
	Castilla-León	0	0	1	2,625.8	54
United Kingdom						
SW	South West	130	0	8	4,666.6	96
SE	South East	703	13	46	17,548	121
EA	East Anglia	277	6	18	2,059.0	100
EM	East Midlands	58	4	12	4,018.8	96
WM	West Midlands	600	7	21	5,219.2	93
NW	North West	113	1	7	6,388.6	95
YO	Yorkshire	104	5	13	4,951.8	93
NO	North	55	1	7	3,075.4	90
GBI	Northern Ireland	42	0	1	1,589.4	78
GBS	Scotland	45	2	8	5,102.4	96
GBW	Wales	45	0	3	2,881.4	87

^a Average for all the country

Table 10: Eigenvalues of the Correlation Matrix of the original data

	Eigenvalue	Cumulative Proportion
First Principal Component	2.70	0.38
Second Principal Component	1.47	0.60
Third Principal Component	1.10	0.75
Fourth Principal Component	0.63	0.84
Fifth Principal Component	0.47	0.91
Sixth Principal Component	0.44	0.97
Seventh Principal Component	0.18	1.00

C Multivariate analysis techniques to define and visualize the clusters of networks

The objective is to define clusters that have small within variation and high between variation. There is a problem associated with the fact that our variables have different characteristics: some variables are continuous (the cost of the network), some are discrete (the size), while others are categorical (the measure of quality) or binary (the dummy for firms as main contractors). To solve this problem, we first extract principal components from the original data set, then group the data based on the principal component combinations. Since we arrive to four principal components we cannot visualize the data directly. We use a canonical discriminant analysis representation to represent the clusters in two dimensions.

The first step consists of extracting principal components from the original data. In Table 10 the eigenvalues are shown. We choose the first four principal components since these explain more 80 % of the variation of the data set.

The eigenvectors associated with these eigenvalues show how the original variables are combined. The eigenvectors are shown in Table 11. The first four principal components associate the variables as follows: The first principal component combines basically the proportion of firms within the network, the dummy for a firm as a main contractor and the fraction (negatively) financed by the Commission. It can be defined as an index of privateness of the network. The second principal component combines the size (negatively)

Table 11: Eigenvectors

Variable	First principal component	Second principal component	Third principal component	Fourth principal component
Size of the network	0.14	-0.40	0.69	0.45
Proportion of firms within the network	0.52	-0.16	-0.05	-0.35
Dummy for firms as main contractors	0.47	-0.15	-0.05	-0.46
Quality of the main contractor	0.21	0.62	0.33	0.02
Fraction financed by the Commission	0.28	0.60	0.19	0.08
Average Quality	-0.51	0.20	-0.07	-0.20
Cost of the project	0.32	0.06	-0.61	0.64

and the measures of quality of the network, hence it can be labeled as an index of quality of the network. The third principal component is a measure of the size and the cost of the network, while the last is a combination of the first and the third. The principal components are the four continuous variables that we will use to cluster the networks together.

We use Ward's (1963) minimum variance method, where the distance between clusters C_K and C_L is defined by

$$D_{KL} = \frac{||\bar{x}_K - \bar{x}_L||^2}{1/n_K + 1/n_L}.$$

In this method the distance between two clusters is based on the Analysis of Variance sum of squares between the clusters. The within-cluster sum of squares is minimized over all partitions obtainable by merging two clusters from the previous set of clusters.

In Table 12 we present the resulting iterative procedure starting with a number of 20 clusters.

Finally, in Figure 1 we visualize the data using canonical discriminant analysis. This procedure just allows to reduce the dimensionality of the problem, in this case four, to a two-dimensional representation. The clusters 1 to 5 show a reasonable small overlap.

D Correspondence analysis to visualize the association of the relevant variables and the clusters

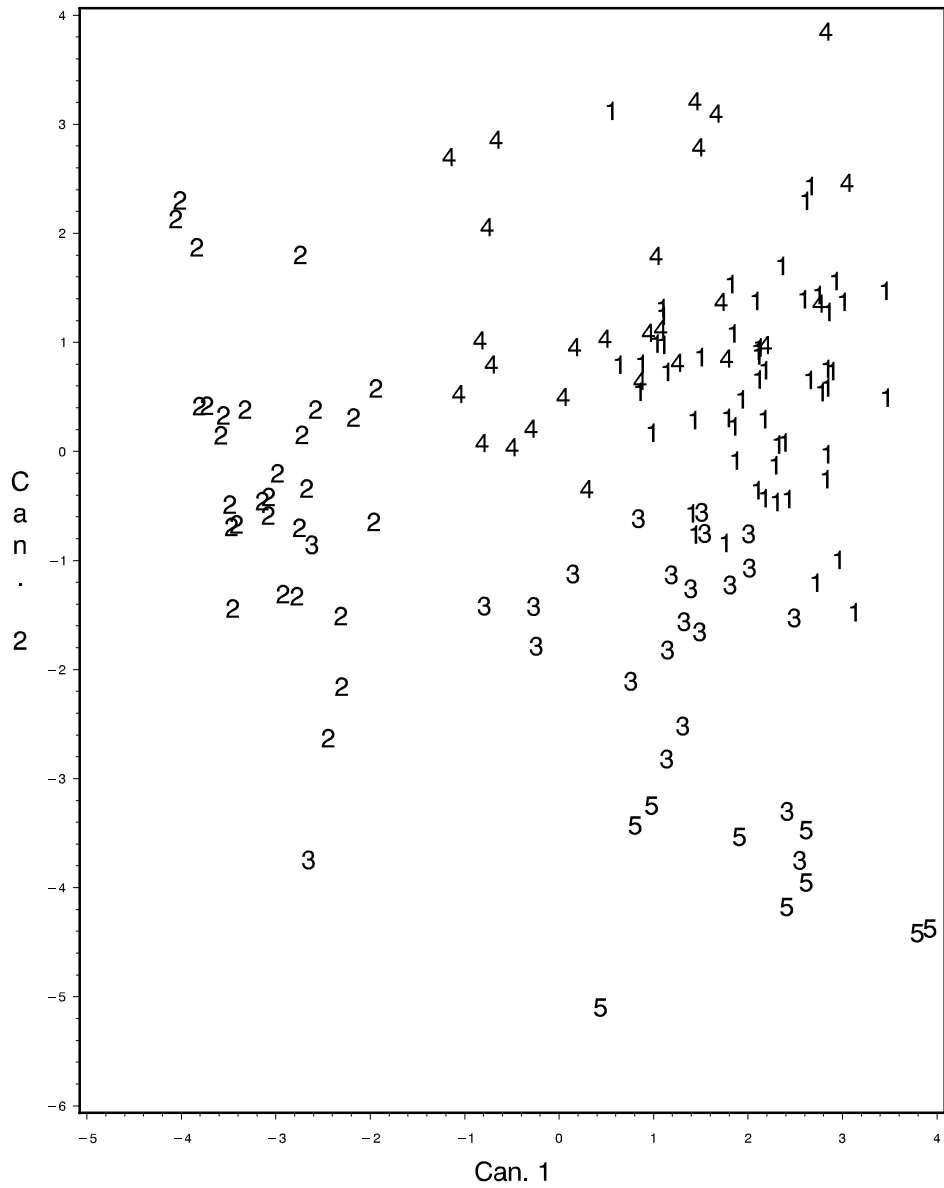
Correspondence analysis is a multivariate technique useful to visualize the association between variables represented in contingency tables. We can think of the clusters as defining a contingency table for each relevant variable. The contingency tables are tables 2 through 7.

The principal output of correspondence analysis is a map where both the row variable, in this case the clusters, and the column variable, in this case variables related to the quality, the type, the size and cost of the networks, can be represented simultaneously. Association can be visualized by the corresponding deviation from the center of the map. For instance in Figure 2 it can be seen that Southern is the cluster more far away from the center of

Table 12: Iterative clustering procedure

Number of Clusters	Cubic Clustering Criterium
20	3.66
19	3.39
18	3.02
17	2.70
16	2.38
15	1.99
14	1.63
13	1.32
12	0.97
11	0.80
10	0.76
9	0.61
8	0.18
7	-0.97
6	-2.08
5	-2.74
4	-1.25
3	-0.40
2	-0.68
1	0.00

Figure 1:
Representation of the clusters



the map. In the same direction of deviation, given by the vertical axis, it can be seen that main contractors of quality groups 1, 2, 3 and 4 are associated with the Southern cluster.

For each map the inertia is reported, which shows how spread the data is across the map, and the quality of the map, which shows the explained variation of the data.

For the details on this procedure see Greenacre (1994).

Figure 2:

Quality of Main Contractor

Inertia = 0.84

Quality = 96.29

Clusters: o

Main Contractor: x

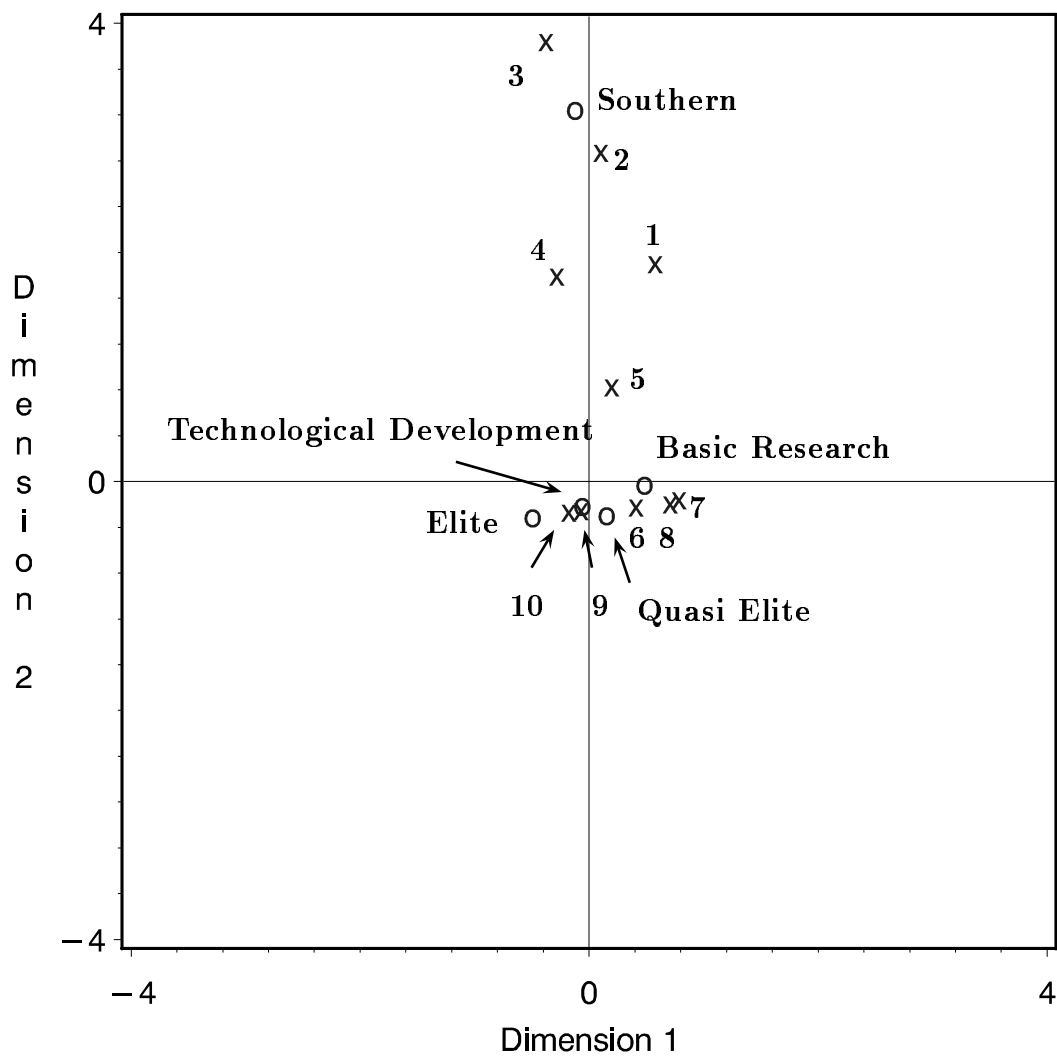


Figure 3:

Average Quality of Network Clusters

Inertia = 0.52

Quality = 88.69

Clusters: o

Average Quality: x

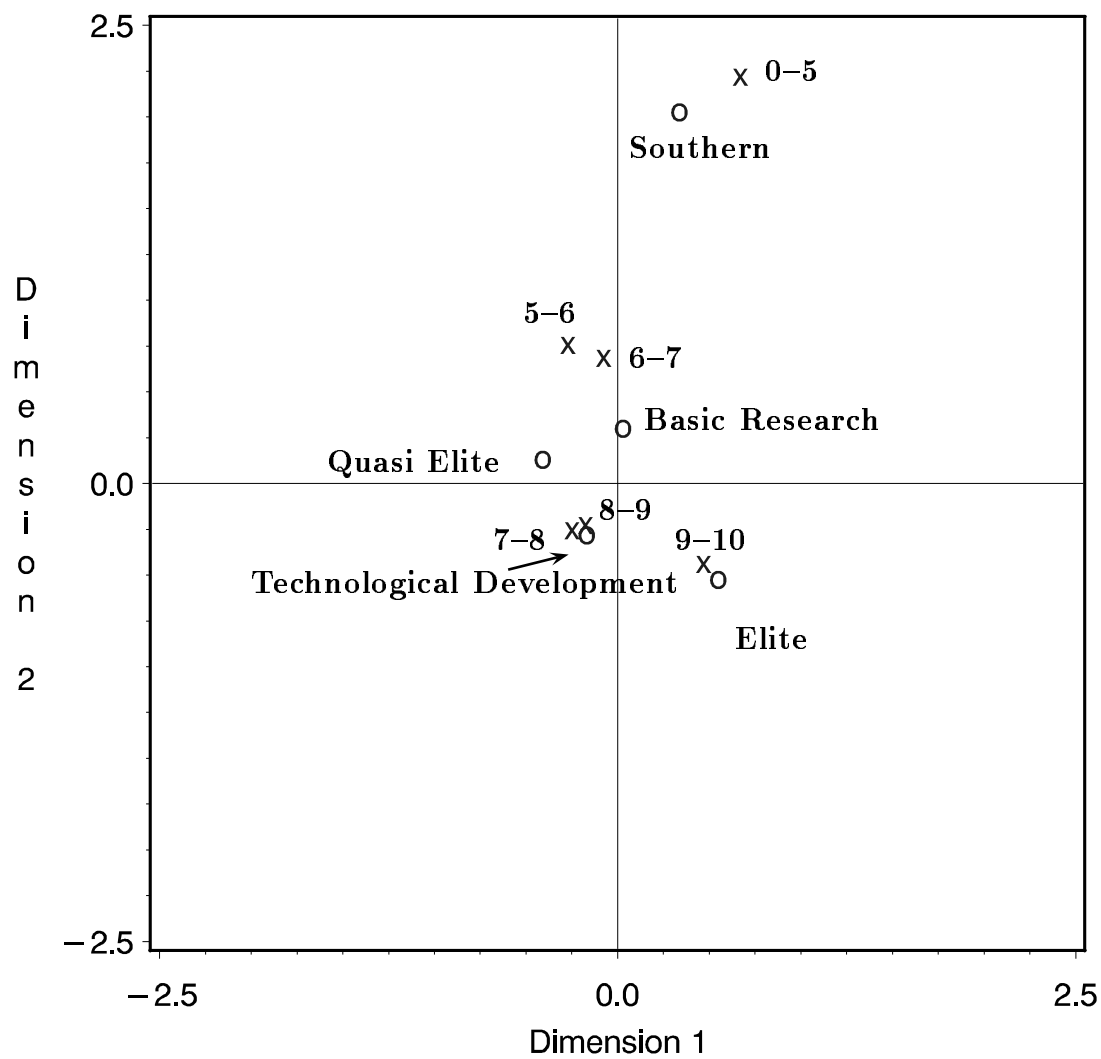
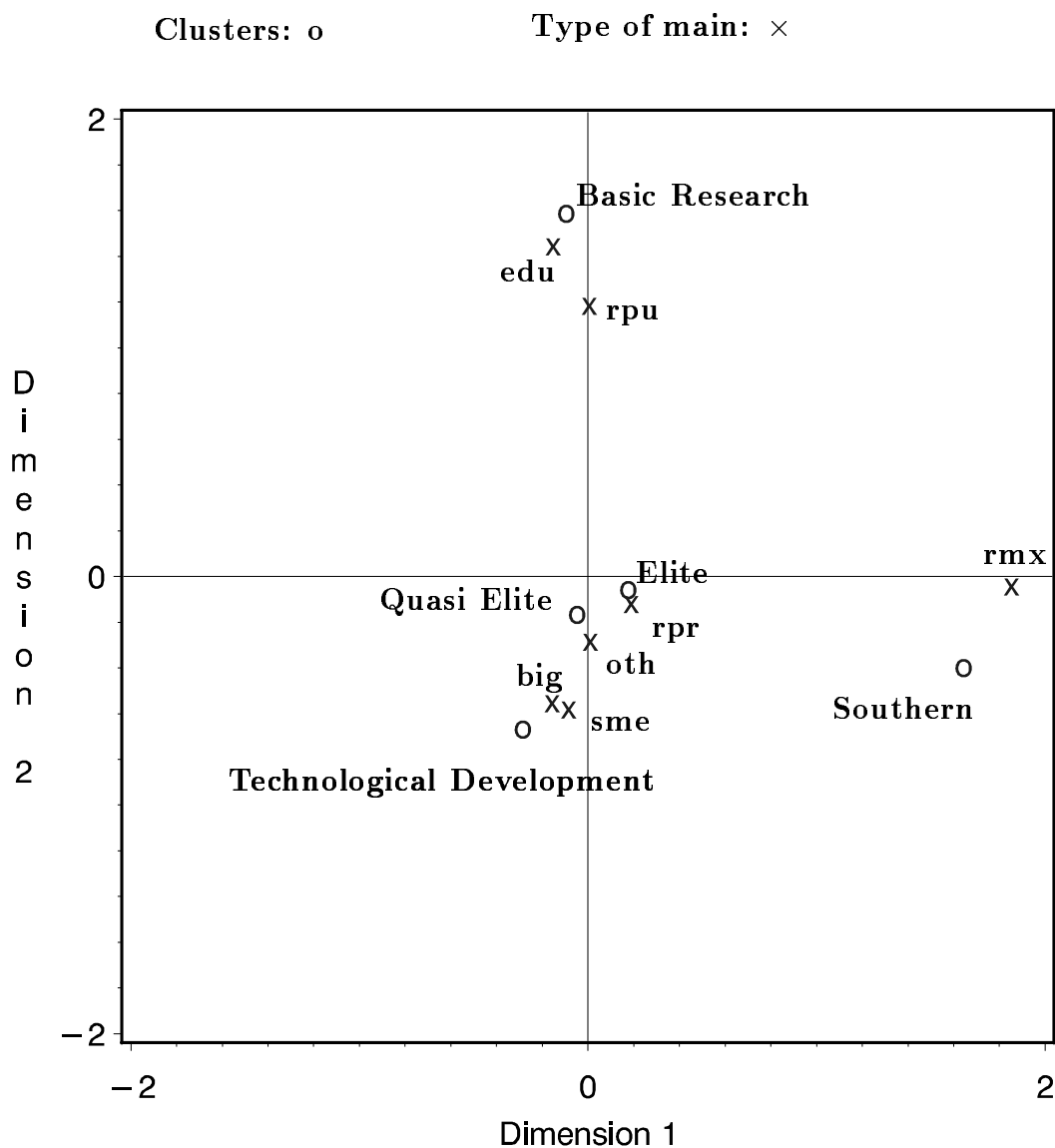


Figure 4:

Type of Main Contractor

Inertia=1.04

Quality=87.08



Big Firms: big, Small-Medium Firms: sme, Universities: edu

Public Research Units: rpr, rmx, rpu, Others: oth

Figure 5:

Proportion of Firms within the Network

Inertia = 1.15

Quality = 93.44

Clusters: o

Proportion: x

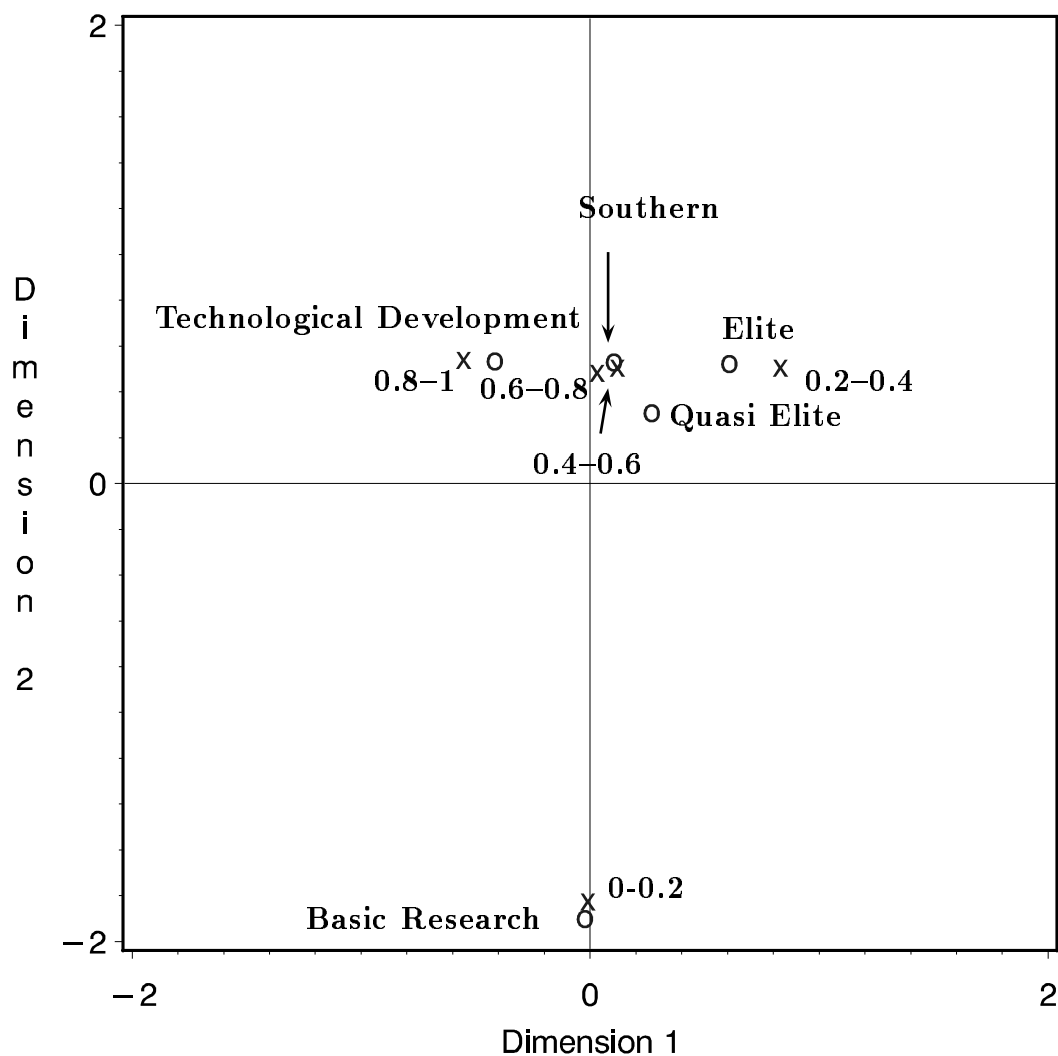


Figure 6:

Size of the Network

Inertia = 0.67

Quality = 91.73

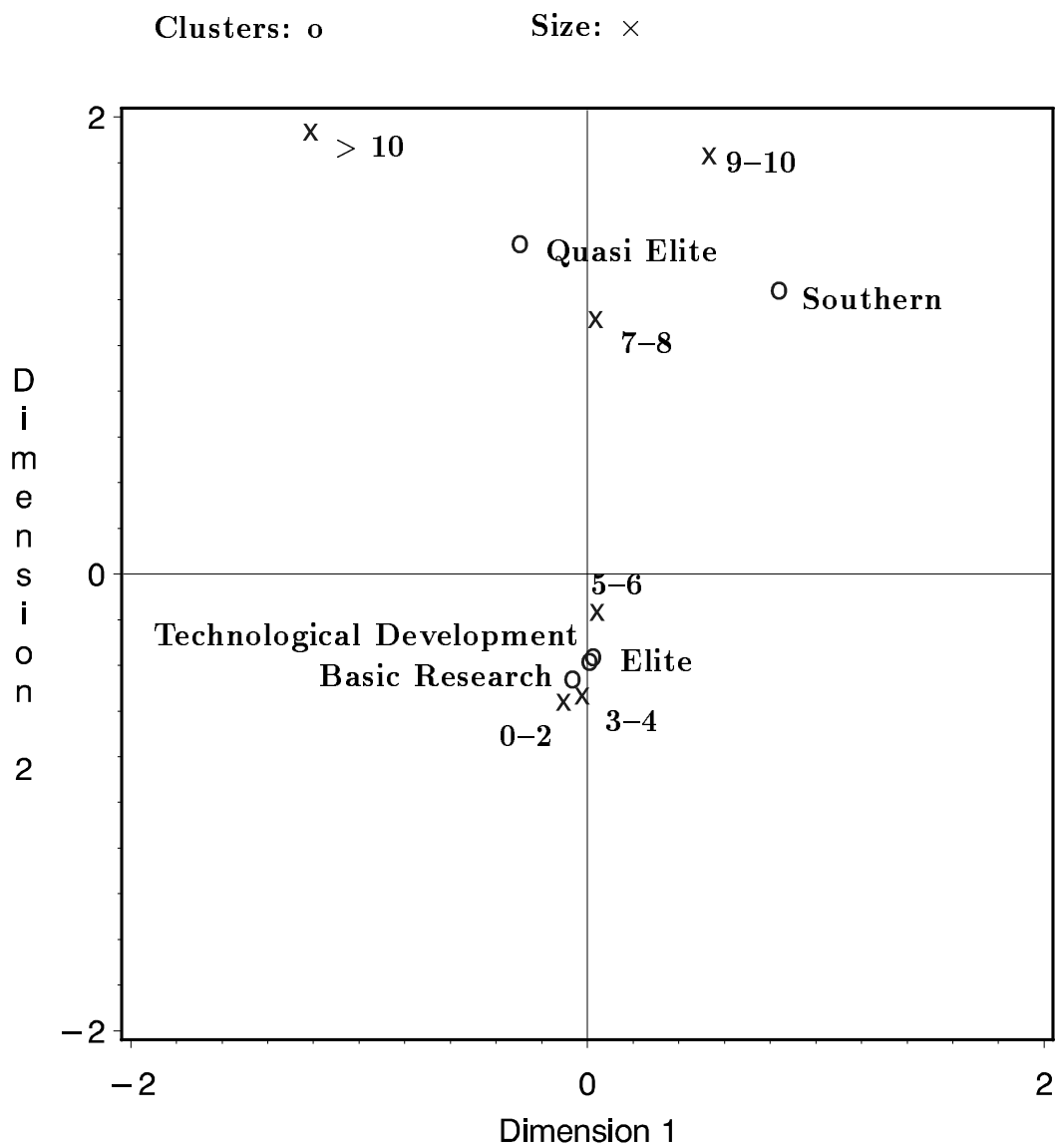


Figure 7:

Cost of the Project (Thousand Ecus)

Inertia = 0.52

Quality = 95.74

Clusters: o

Cost: x

