



Firm Segmentation as a Tool for R&D Policy Evaluation: Revisiting the Taxonomy of Firms Engaged in International R&D Networks

Bruno Brandão Fischer¹, José Molero²

Abstract

As we have addressed in an earlier article, quantitative evaluation of R&D policies often approaches the outcomes of initiatives without effectively considering differential impacts on economic agents. Our goal is to confirm the existence of “segments” of firms according to their outcomes arising from participation in a given R&D program. Data is gathered from Eureka Program’s Final Reports (2000-2005 and 2006-2008) from Spanish, Italian, French, British, and German firms. Classification focus is directed towards variables related to innovation projects’ outcomes. Log-likelihood clustering method was used. Tests for differences between clusters in terms of some variables of interest were performed. Results are consistent with the hypothesis of marked heterogeneity in firms’ outcomes. This methodology offers a valuable instrument for RTD policymakers in terms of monitoring and ex post intervention.

Keywords: r&d policy; evaluation; eureka program; clusters; european union.

Universidad Complutense de Madrid/Research Group on Economics and Innovation Policy. Campus de Somosaguas-Finca Mas Ferré. Edificio A. 28223 – Madrid, Spain. e-mail: ¹bbfischer@pdi.ucm.es, ²jmolero@icei.ucm.es

Introduction

Innovation policy is about generating desirable behavior in firms through the creation of a framework of incentives for the supply-side of the economy. According to a Schumpeterian view of economic systems, innovation would generate overall growth through the establishment of new and improved industries and business practices. We then part from the assumption that innovation is a central feature of the process of economic growth and development (Metcalfe, 1995). Therefore, related policies are a matter of great concern worldwide where the science-technology-innovation system is continuously and rapidly evolving and integrates a context in which business competition is increasingly based on its terms (Freeman and Soete, 2009).

In the European Union this situation is not different, even though some idiosyncrasies of it make this bloc an interesting case for studies. The process of integration leads to incentives for agents to further invest in innovation, since access to a larger market becomes readily available. Europe needs to be more innovative in order to achieve a higher level of competitiveness – which requires a change in the valuation of innovative activities and a better framework for innovation to develop (European Commission, 2006). A direct implication of this perspective, especially in times of crisis, is that RTD investments need to become more productive, considering that there is not enough capacity for larger amounts of funding.

In operational terms, this means that the evaluations of such initiatives happen to be extremely important. If more efficient application of resources is required, then characteristics of R&D expenditures must be understood deeply (Benfratello and Sembenelli, 2002). Nonetheless, the process of RTD policies' evaluation is somewhat multidimensional and it presents blurry connections in terms of causality and its inherent direction (bi-directional relationships are often found in literature), thus making it difficult to evaluate the actual impact of these activities (see, for example, García, 2010). We can attribute this point of view to the fact that innovation involves complex interactions between agents and the environment they are embedded in (Smith, 2000). This gives an idea of the complexity involved not only in formulating innovation policies, but also in evaluating their impact, and that is why a proper management of innovation is a challenge for many countries, including developed ones (Aghion and Tirole, 1994).

The foundation of such assessment lies on the authors' uneasiness with the relative lack of literature focusing on differential impacts of innovation policies on companies. While there are extensive studies of policies' effectiveness based on quasi-experimental methods as well as on impacts on

companies' innovative and/or economic performance, we find a gap in terms how these policies affect different companies in heterogeneous ways. If we regard evaluation of RTD policies as a linear phenomenon which would impact firms homogeneously, implausible assumptions would have to be accepted, where overall results could easily be distorted by influent observations, driving evaluation results positively or negatively while not being representative of actual developments (Fischer, 2012).

Companies have long acknowledged that addressing diverse audiences with a uniform supply would result in a large proportion of unsatisfied customers. Strategies of market niches and segmentation were developed, where subpopulations could be approached in an (almost) optimal manner. Similarly, firms represent the "market" for innovation policy, but how can standardized initiatives generate adequate benefits for agents that are distinct in so many dimensions (size, R&D intensity, short and long term objectives, etc.)? Consequently, how can the evaluation of RTD policies perform fittingly if it is undertaken as if firms interact with such programs in a similar manner?

Our expectation is to build upon these findings a useful methodology to classify firms participating in international R&D cooperation projects and to extract information regarding patterns of technological and innovative behavior according to companies' outcomes resulting from their participation in a given project. This rationale extends our approach applied to case of Spanish companies (Fischer and Molero, 2012), considering a broader range of countries (German, French, British and Italian firms were added to Spanish ones) and further statistical tests were undertaken in order to analyze differences between resulting groups. Such analysis should provide policymakers with relevant managerial information regarding processes of international R&D collaboration between agents, as well as a suitable tool for innovation programs' monitoring.

The paper is structured as follows: chapter 2 depicts aspects related the economics of RTD policies, as well as brief notes on their evaluations. Chapter 3 develops on the methodological approach, stating main features of the case under scrutiny (Eureka Program), the data and the clustering process, as well as statistical tests. Chapter 4 reports the outcomes and compares them with our previous findings. Chapter 5 offers some final remarks and policy implications of our assessment.

Economic Rationale Behind Rtd Policies

Technological innovation policies represent a strategic area in the field of public policy regardless of governments' politi-

cal inclination or geographical relevance (national, regional, local or even supranational) (Aghion and Tirole, 1994). This is a result of the role that innovation and technological change play in fostering economic growth and its characteristics of public goods that are likely to create market failures (Branstetter and Sakakibara, 2002; Suurna and Kattel, 2010). However, RTD policies need to evolve, along with a changing socioeconomic environment, as well as regarding issues that leave room for improvement. Thus, they are often evaluated, adapted and modified in order to provide society with better outcomes, and agents with a better framework of action. Adaptive policymaking is about facilitation (enabling innovation), understanding the existence of unpredictability and indeterminacy in the results of policy initiatives (Metcalf and Georghiou, 1997). Furthermore, innovation processes happen in conditions of uncertainty and (in the capitalist system) of competition and so must be approached in a holistic manner, considering not only technical capabilities but also the market environment and the social context (Pavitt, 2003; Kline and Rosenberg, 1986).

Innovation is a costly process which can create market failures related to the nature of technological change, such as appropriability issues, amount of R&D investment, spillovers and externalities (Klette and Moen, 1999; Smith, 2000; European Commission, 2006a; NIST, 2006). Arrow (1962) argues that innovation's inherent characteristics demand governmental action in order for society to produce an optimal level of economic valuable knowledge. This rationale of "market failure" still is predominant in justifying the need for public policies that approach the problems related to the innovative process, fostering an innovation-driven environment (Bayona-Sáez and García-Marco, 2010; Nelson, 1959; Sanz Menéndez, 1995). The main argument is that in a context of perfect competition, there are not enough incentives for firms to innovate, given the lack of economic institutions that guarantee the return on investments. As examples of policies based on the "market failure" mindset we can highlight Intellectual Property Rights and R&D subsidies.

Nonetheless, it is known that the mere understanding of market failures does not provide enough information for technological policymaking (Aghion, David and Foray, 2009). Even though neoclassical approaches have been widely recognized as useful for the RTD policymaking process, the evolutionary theory in economics has also contributed in this arena with a systemic orientation, mainly through the framework contained in the National Innovation Systems framework of analysis. The Innovation System approach considers the economic and social environment for innovation as one where agents do not innovate in isolation, but rather through complex interactions.

Despite its differences, innovation policy, either neoclassical or systemic/evolutionary, involves industrial, environmental, labor and social aspects, aiming at the generation of economic competitiveness (Kuhlman and Edler, 2003). But evolutionary theory influenced technological policies to become more oriented to adaptation of firms and markets in an environment of change (Nelson and Winter, 2002). They provide the framework for understanding the own system's changes over time. We can affirm then that existing institutional structures, including bodies of relevant law, and particular government policies and programs, can never be regarded as optimal and for this reason they are, and should be, always subject to evaluations and constructive criticism (Nelson, 2007).

During the last decades, globalization and the shift towards knowledge as the source of competitiveness rendered the traditional policy instruments less effective (Gilbert, Audretsch and McDougall, 2004), creating an environment that demands continuous adaptation in public policies and initiatives (Bin and Salles-Filho, 2012). Technology policies are part of a complex economic landscape and must ensure that firms are able to realize their innovative potential, meaning that the appropriate R&D policymaking requires knowledge about context conditions, group behavior, instruments (and their mix) and policy effects (Ebersberger, Edler and Lo, 2006).

Referring to this latter argument, given the complexity of business environments and different sectoral characteristics, innovation policies cannot afford to be fully standardized, since there is no optimal design for them: these vary across countries, technological domains and stages of innovative processes (Raymond and St. Pierre, 2010; Klette and Moen, 1999). Provided there is a high level of complexity and dynamism in the policymaking process regarding innovative activities, as well as the need for adaptation at the innovation system level, evaluation activities become a key element in designing better programs to develop innovative capabilities and desired changes in behavior and structure.

Evaluation of Rtd Policies

Since R&D policies can be considered fundamental for long-term development and are subject to an ever-changing environment, there is a strong need to continuously evaluate their effectiveness (Bayona-Sáez and García-Marco, 2010). Furthermore, emphasis should be given to policy trials and their evaluation: the process of adaptation may consist in trials and errors (Metcalf and Georghiou, 1997). There is a continuous need for better understanding of innovation processes and policies aiming at its promotion (European Commission, 2002), especially because innovation is disruptive by nature, and it breaks established patterns of behavior,

giving rise to unpredictable consequences” (Metcalf, 1995). This assertion brings to light the fact that investment in new knowledge is not an exact science and will not necessarily provide firms with the anticipated returns in terms of competitiveness, which also indicates that these investments may not turn into commercialization of outcomes (Audretsch and Keilbach, 2008). Nonetheless, it is fundamental for market-oriented innovation policies to take into account not only technical aspects, but also potential and actual market impacts of projects (NIST, 2006).

In order to cope with such scenario, technological policy evaluation provides a systematic and valuable way of adaptive learning based on the analysis of practical situations, thus representing a resource of great potential for policymakers (Georghiou, 2002; Malik and Cunningham, 2006; NIST, 2003). The process of analyzing and evaluating RTD policies represents means to improve the policymaking process, both in terms of policies’ suitability to a specific context and to achieve managerial progresses in existing programs. The structure of relationships within a system, knowledge flows, existing capabilities and market conditions also shape the scenario for an innovative environment to develop. This poses the relevance of innovation policy in fostering a mosaic of desirable characteristics and that might take some time to be implemented in a society, which can only be accomplished through methodical and frequent evaluation (European Commission, 2006a).

Evaluation provides measures for success, thus contributing to evolution and improvements of existing initiatives (NIST, 2003), which is dealt mainly with assessments on programs’ efficiency and efficacy, i.e., how well the initiative worked, if achieved its goals or not, and its contribution to the overall policymaking scenario (European Commission, 2006a; Georghiou and Keenan, 2006). Hence, evaluation activities consist basically in systematically and objectively determining the relevance, efficiency and effect of an activity considering its goals, providing policymakers with feedbacks on the impacts of such initiatives and creating fundamental knowledge for the promotion of necessary adjustments for future policies’ formulation and implementation (Durieux and Fayl, 1997; European Commission, 2002). Besides making it possible for program managers to assess the benefits of a given initiative, and to identify opportunities for improvement, RTD policy evaluation allows the communication of program’s results to society (US Department of Energy, 2007). Currently, the growing complexity involved in technological generation increase uncertainties on impacts from RTD policies (European Commission, 2002). Moreover, there has been an enlarging trend in terms of policy instruments, thus implying a need of a more diverse and complete group of analytical tools. Accordingly, evaluation activities and the identification of policy “best practices” in OECD countries

represent a complicated task given the myriad of technological initiatives that take place in these nations (ranging from basic research direct support to more indirect measures aimed at improving the capacity of firms to innovate and use new technologies) (Durieux and Fayl, 1997). Like science in general, technological policy evaluation might also be considered as a research and scientific matter (Georghiou, 1999). Based on this context, our criticism to widespread techniques of RTD policy evaluation lies on the use of methods that analyze a large sample of firms as if they behave or interacted homogeneously with a given initiative on the one hand, and those methods that base their analysis on case studies, providing results that are hardly able to provide insights that are applicable to other firms. These are extremes that do not comply with practical possibilities of RTD policy evaluation, as well as do not take advantage of information available. In the next section we outline a simple method to combine the utility of large samples and the idea of “customization” of evaluations, trying to gather the benefits of both kinds of assessments.

Methodological Approach

Data for this research comes exclusively from Eureka individual projects’ dataset of final reports, which was provided by the Eureka Secretariat. Such reports are structured as questionnaires, containing several questions on different aspects. The Eureka Programme was created aiming at enhancing collaboration between companies in a market oriented, non-bureaucratic, bottom-up approach promoting cooperative projects for national funding (Georghiou, 2001; Marín and Siotis, 2008). Eureka is present in 38 countries and does not act through financial support, but providing projects with a seal of approval that facilitates access to governmental funds in the national level (Georghiou and Roessner, 2000). Its focus is on improving European competitiveness and productivity through an enhanced cooperation between companies from different Member Countries (international collaboration) and research centers in high-tech areas.

The timeframe used is based on two different periods: 2000-2005 and 2006-2008 (dates of projects’ conclusion). These datasets are analyzed separately because of operational issues. The instrument of data collection suffered changes in between these periods, altering aspects such as the existence of certain variables of interest, as well different sorts of measurement scales. In geographical terms, we assessed data for five European countries: Spain, Italy, UK, France and Germany. These can be regarded as highly representative of the European situation, gathering data for the largest economies and which face different stages of development in terms of their innovation systems. It is important to notice that such observations referred to finished projects, while our analyses used firm-level data. This is justified by the fact that by

using data of projects we would inflate the influence of variables related to companies, since several companies were involved in more than one project. Thus, we merged data for such companies. The resulting structures of datasets were the following:

a) 2000-2005 - 77 Spanish firms; 60 German firms; 34 French firms; 27 Italian firms; 17 British firms. N = 215.

b) 2006-2008 - 36 Spanish firms; 52 German firms; 19 French firms; 6 Italian firms; 2 British firms. N = 115.

Based on the outcome aspects of projects – TECHACHIEV, COMMACHIEV and EXP_IMPACT (table 1) - we proceeded to a categorization of the firms being analyzed for both periods under study (2000-2005 and 2006-2008). For this approach, the TwoStep Cluster (SPSS) method was used. This method is an exploratory tool designed to reveal natural clusters in the dataset according to the parameters indicated, offering the possibility of suggesting latent taxonomies. Also, we used Log-likelihood distances to build the clusters, since this procedure allows the use of categorical variables, which is not possible with Euclidean estimations.

To establish the optimal number of clusters we developed on the structure proposed by Fischer and Molero (2012). Thus, we shall test for the consistency of 3 clusters. These authors proposed the following categorization of firms in an exploratory assessment of Spanish firms in Eureka 2000-2005:

(i) Risky Innovators - companies comprehended in this cluster had the best technical outcomes out of the three clusters, but only partially they can obtain satisfactory market results.

(ii) Inventors - These companies were classified as inventors for showing fair technical results without taking advantage of it in the market – which does not allow us to define them as innovators per se.

(iii) Consistent Innovators - These companies had poorer technical results than risky innovators, but they consistently achieve positive commercial results.

This procedure aimed at generating clusters of firms from all countries included (Spain, Germany, France, UK, and Italy), checking for robustness over time through the comparison of cluster structure between 2000-2005 and 2006-2008 datasets. The goal here is to provide a consistent perspective of behavioral patterns of agents, according to their projects' outcomes.

Furthermore, we tested for differences between clusters in terms of some variables of interest, namely (for definitions see table 2): TOT₁_COST, DURATION (via parametric tests, i.e., ANOVA), RATIO_RD, FUNCTIONING, and COUNTRY (via non-parametric tests, using Mann-Whitney U statistics, provided that these variables are ordinal). This assessment allows a deeper comprehension of clusters, as well as a complementary approach to relationships between

| Code | Description | Structure | |
|-------------|--|---|--|
| | | 2000-2005 | 2006-2008 |
| TECHACHIEV* | Evaluation of Overall technological achievements in the project. Source: Eureka | 1 = Excellent 2 = Good 3 = Weak 4 = Bad | 1 = Excellent 2 = Good 3 = Weak 4 = Bad |
| COMMACHIEV* | Evaluation of commercial achievements as a results of the project. Source: Eureka | 1 = Excellent 2 = Good 3 = Weak 4 = Bad 5 = Nil | 3 = 2-5 times RTD investment 4 = +10 times RTD inv. 5 = <100 times RTD inv. 6 = More 1 = Very low 2 = < or = RTD investment |
| EXP_IMPACT* | Expected future impact of results from the project. Source: Eureka | 1 = Very Large 2 = Large 3 = Medium 4 = Small 5 = Nil | 3 = 2-5 times RTD investment 4 = +10 times RTD inv. 5 = <100 times RTD inv. 6 = More 1 = Very low 2 = < or = RTD investment |

Table 1. Outcome variables' structure for cluster analysis

variables that can indicate some form of association which can be regarded as an influential cause of success in international R&D cooperation projects.

* Methodological note I: Countries' codes are assigned according to their relative position in terms of the stage of development of their National Innovation Systems. The higher the rank, the more developed. Furthermore, they were grouped in three categories, where Spain is referred to as a laggard Innovation System; Italy, UK and France are classified as intermediate Innovation Systems (including, thus, lower intermediate, Italy, intermediate, UK, and upper intermediate, France); and Germany is regarded as the leading nation in terms of IS capacities.

** Methodological note II: Categories 3 and 4 (Weak and Bad Functioning, respectively) were merged in order to have analyzable data, since observations in category 4 were scarce.

The use of clustering algorithms can be regarded as an interesting tool for policy monitoring and post hoc evaluation, since it offers a dynamic view of the interaction between influential variables in the determination of agents' relative position in the process, considering that such clusters present a stable structure.

Cluster Analyses

This section contains two estimations of clusters. The first one represents clusters for 2000-2005 firms according to their outcomes (technological, commercial, and expected). These results are confronted with those found by Fischer and Molero (2012)* in an exploratory assessment of Eureka's projects in Spanish firms. Their proposed taxonomy is applied to clustering structures, and similarities, as well as differences, are discussed. In order to further exploit the usefulness of achieved clusters we test for differences between groups according to variables which correspond to Microeconomic (RATIO_RD), Contextual (FUNCTIONING, TOT_COST, DURATION), and Macroeconomic (COUNTRY) dimensions. For this, ANOVA estimations and non-parametric (Mann-Whitney) statistics are applied. A second stage contains a similar assessment for 2006-2008 data.

*Fischer and Molero (2011) use a dataset consisting of Spanish companies only. However, criteria for groups are refined in this present work, since focus is given on firms' outcomes. Evaluation between groups characteristics regarding other aspects of microeconomic, contextual, and macroeconomic dimensions are tested statistically through parametric (ANOVA) and non-parametric (Mann-Whitney) approaches. This procedure makes possible a stronger internal consistency within clusters, while statistical tests provide more robust perceptions on their differences regarding other variables of interest.

| Code | Description | Structure | |
|---------------|---|---|---|
| | | 2000-2005 | 2006-2008 |
| TOT_COST | Total cost of project(s) carried out by firms. Source: Eureka | Millions of euros | Millions of euros |
| DURATION | Duration of project(s). Source: Eureka. | Months | Months |
| NIS* | Consists of countries' categories to which firms belong, i.e., Spain, Germany, France, UK, and Italy. It functions as a proxy for National Innovation Systems' characteristics. | 1 = Spain 2 = Italy; United Kingdom; France (Intermediate) 3 = Germany (Leader) | 1 = Spain 2 = Italy; United Kingdom; France (Intermediate) 3 = Germany (Leader) |
| RATIO_RD | Ratio between R&D expenditure and total turnover. Source: Eureka | 1 = <2% 2 = 2 to 10% 3 = >10% | 1 = <2% 2 = 2 to 10% 3 = >10% |
| FUNCTIONING** | Evaluation of functioning's quality of project's participants. Source: Eureka | 1 = Excellent 2 = Good 3 = Weak/Bad | Not included |

Table 2. Control variables for cross-cluster comparison

2000-2005 Segmentation

This first part of cluster analyses considers results for companies with Eureka projects ending within the period 2000-2005. Results regarding clusters' structures are provided in table 3. Distribution of cases is mainly concentrated in cluster 1, while clusters 2 and 3 are of similar sizes. Variables used in this analysis turned out to be statistically significant separators of the three obtained groups. In a first moment, we verify clusters' features in order to observe the main trends in terms of outcomes (technological, commercial, and expected) and then we confront these findings with the taxonomy proposed by Fischer and Molero (2012).

Cluster 1 has a strong orientation towards fair (but not exceptional) achievements. This cluster is characterized technologically by firms achieving Good results. Furthermore, commercial achievements are positive, even though there is a relevant presence of firms with Weak attainments in this

regard. Lastly, expected impacts for cluster 1 are also positive, but not outstanding, where Medium and Small impacts are predominant. Cluster 1, thus, can be pointed as analogous to the Consistent Innovators group, where firms have good technological results (but poorer to those attained by cluster 3, as we shall discuss below) and consistent market outcomes (achieved and expected). Such firms represent the bulk of this particular dataset, where innovation is likely to take place, and competitiveness of companies shall be maintained (or even increased). Nonetheless, it is not likely that such outcomes will have the broad impact that can be expected from groundbreaking innovations. We find support to this result in Georghiou's (2001) criticism on the declining impacts of Eureka projects in general, where risks are not being fully taken by agents.

Cluster 2 is similar to cluster 1 in terms of technological results, where they are largely rated as Good. Nonetheless, perceptions on market impacts point in a different direc-

| Cluster Distribution | | Correspondence to Fischer and Molero's (2012) Classification | |
|----------------------------------|--|--|--|
| | Cluster 1 | 116 (54.2%) | Consistent Innovators |
| | Cluster 2 | 46 (21.5%) | Inventors |
| | Cluster 3 | 52 (24.3%) | Risky Innovators |
| | Missing | 1 (.5%) | |
| Cluster Profiles and Main Trends | | | |
| | Cluster 1 | Cluster 2 | Cluster 3 |
| TECHACHIEV | Good Technological Achievements (94%)* | Good Technological Achievements (71.7%)* | Excellent Technological Achievements (98.1%)* |
| COMMACHIEV | Good (58.6%), and Weak (36.2%) Commercial Achievements* | Nil Commercial Achievements (89.1%)* | Excellent (26.9%), and Good (48.1%) Commercial Achievements. Also, 13.5% of Nil Commercial Achievements* |
| EXP_IMPACT | Large (22.4%), Medium (42.2%), and Small (30.2%) Expected Impacts* | Medium (19.6%), Small (32.6%), and Nil (45.7%) Expected Impacts* | Very Large (19.2%), Large (32.7%), and Medium (40.4%) Expected Impacts* |

*Clusterwise Importance (chi-square at 95% confid.)
Percentages correspond to within cluster data.

Table 3. 2000-2005 Clusters

tion, where Nil Commercial achievements are predominant. If this could indicate that such firms may have a longer time-to-market period, the analysis of EXP_IMPACT shows that in reality these agents also do not expect further results to unfold over time. The relative position of cluster 2 is the worst out of the three clusters in terms of market appropriation of results (and expectations regarding future outcomes). There is a strong correspondence between this particular group of firms and the Inventors classification, where the technological part of projects is fair, but market results are disappointing, indicating that little economic impact arose from participation in an international R&D cooperation project. Nonetheless, such activity might have played a role in enhancing firms' absorptive capacities and technological capabilities, thus contributing to its overall performance in structural terms.

Cluster 3 shows a strong participation of firms with Excellent technological results, thus rating this particular cluster in a better position than the other two. In commercial terms, this cluster also contains companies with the best results, even though stronger variability exists in this regard that is not observed in the other clusters. For cluster 3, Excel-

lent commercial results are reported, but also Nil commercial results. Nonetheless, future results are expected to be better than those perceived by respondents of the other groups. The correspondence of this cluster with Fischer and Molero's (2012) taxonomy finds no perfect match, but the presence of an unexpected variability in commercial achievements expresses the existence of riskier projects, where achievements are more relevant, but also are not necessarily associated with positive market outcomes. This leads us to classify this cluster as Risky Innovators.

If we confront these three clusters regarding some relevant variables for the evaluation of the hypotheses formulated in this research, further insights can be found in addition to those provided by logistic regressions. A summary of such approaches is reported in table 4. Categorical variables were tested through Mann-Whitney non-parametric tests for independent samples, where continuous variables were assessed through ANOVA (Bonferroni and Tamhane's post hoc tests were assigned according to results of Levene's homogeneity of variance tests). Also, descriptive statistics are provided for each variable within each cluster.

The variable *RATIO_RD*, a measure of firms' innovative intensity and a proxy of absorptive capacity, does not result in an efficient indicator of differences between clusters. This can be gathered from an evaluation of means, medians, and standard deviations. The Mann-Whitney test for non-parametric differences between samples confirms this perception, where cluster comparisons are not significant in any pairwise contrast. This finding is especially relevant for understanding the non-linear behavior of technologi-

| <i>Means (Medians in parentheses; Std. Dev. in brackets)</i> | | | |
|--|-----------------------|-----------------------|-----------------------|
| | Cluster 1 | Cluster 2 | Cluster 3 |
| RATIO_RD | 2.12 (2) [.677] | 2.05 (2) [.631] | 2.05 (2) [.714] |
| FUNCTIONING | 1.97 (2) [.407] | 1.96 (2) [.631] | 1.29 (1) [.502] |
| COUNTRY | 1.93 (2) [.805] | 1.84 (2) [.759] | 1.94 (2) [.826] |
| Non-Parametric Tests for Independent Samples | | | |
| Mann-Whitney Tests for Pairwise Comparison | | | |
| | Cluster I | Cluster J | Sig. |
| RATIO_RD | 1 | 2 | .532 |
| | 1 | 3 | .570 |
| | 2 | 1 | .532 |
| | 2 | 3 | .992 |
| | 3 | 1 | .570 |
| | 3 | 2 | .992 |
| FUNCTIONING | 1 | 2 | .799 |
| | 1 | 3 | .000 |
| | 2 | 1 | .799 |
| | 2 | 3 | .000 |
| | 3 | 1 | .000 |
| | 3 | 2 | .000 |
| COUNTRY | 1 | 2 | .525 |
| | 1 | 3 | .991 |
| | 2 | 1 | .525 |
| | 2 | 3 | .585 |
| | 3 | 1 | .991 |
| | 3 | 2 | .585 |

| <i>Means (Std. Dev. in parentheses)</i> | | | |
|---|-------------------|-------------------|-------------------|
| | Cluster 1 | Cluster 2 | Cluster 3 |
| TOT_COST | 7.362 (15.709) | 3.383 (3.396) | 2.853 (3.731) |
| DURATION | 44.97 (22.716) | 41.48 (16.591) | 39.10 (21.256) |

| One-Way Anova Results | | | | |
|--|-------------|-------------|---------------------|-------|
| Tamhane and Bonferroni's Post Hoc Tests | | | | |
| | Cluster (I) | Cluster (J) | Mean Difference I-J | Sig. |
| TOT_COST* | 1 | 2 | 3.978 | .032 |
| | 1 | 3 | 4.508 | .012 |
| | 2 | 1 | -3.978 | .032 |
| | 2 | 3 | .530 | .845 |
| | 3 | 1 | -4.508 | .012 |
| | 3 | 2 | -.530 | .845 |
| DURATION** | 1 | 2 | 3.487 | 1.000 |
| | 1 | 3 | 5.869 | .296 |
| | 2 | 1 | -3.487 | 1.000 |
| | 2 | 3 | 2.382 | 1.000 |
| | 3 | 1 | -5.869 | .296 |
| | 3 | 2 | -2.382 | 1.000 |

*Tamhane's post hoc tests. Levene's sig. .001

**Bonferroni's post hoc tests. Levene's sig. .293

Table 4. 2000-2005 Clusters' ANOVA and Mann-Whitney tests

cal innovation. On the one hand, R&D intensity should lead to better technological and market-related results, if one would resort to a linear manner of understanding causality in technological management. On the other hand, this indicator shows that segments of firms that differ in terms of results do not differ substantially in terms of absorptive capacity (considering this variable as a proxy for such). This can be regarded as rather surprising, since results refer to collaborative projects.

A similar outcome is achieved for COUNTRY, indicating that clustering procedures according to firms' outcomes (TECHACHIEV, COMMACHIEV, and EXP_IMPACT) do not produce internally homogeneous groups in terms of this variable. We can gather from this result a marginal role of nationality in shaping firms' ultimate achievements. National Systems of Innovation, or the macroeconomic scenario as a whole, are likely to influence its agents' capabilities and characteristics, but again, not similarly. Segments in this case are weak indicators of firm nationality, regardless of attainments.

However, while the variable FUNCTIONING has a comparable structure between clusters 1 (Consistent Innovators) and 2 (Inventors), cluster 3 (Risky Innovators) is composed by better rates of project functioning (lower ratings represent better results; refer to table 2 for variables' descriptions). We find statistical support for this indication in non-parametric tests, where cluster 3 is relevantly different from clusters 1 and 2. This finding provides support for the hypothesis that better results (technological and economic) are significantly related to projects that excel in managing networks properly. Even though cluster 3 has a stronger variance in terms of market achievements than cluster 1, firms from the former also achieve more relevant economic outcomes, thus excelling in innovative attainments. As the initiative under scrutiny makes reference to collaborative projects that happen in an international context, one can assume that the framework in which innovations develop has its complexity increased by the addition of inter-firm transaction costs. It is somewhat logical, then, to expect a higher relevance for a variable that represents managerial quality.

In the case of the continuous variables, TOT_COST and DURATION, differences between groups show that costlier projects are associated with cluster 1 (Consistent Innovators), which shows satisfactory, but not outstanding, results for firms. This particular variable complements, as pointed out before, the innovative intensity of projects (whereas RATIO_RD represents innovative intensity of firms). The conclusion that can be drawn from this analysis is that projects with a safer return involve larger amounts of investment, even if such investment shall not provide excellent effects. Under a different perspective, firms might be more critical when evaluating the results of these projects. Under this perspective, cluster 3 would not necessarily outperform

cluster 1, since opinions about achievements can be biased by the size of investment made, while firms from cluster 3 could be more sensitive to results in face of smaller R&D expenditures. This brings us to a discussion on the productivity of projects, and self-reports do not allow for a robust conclusion in this case. Nonetheless, such aspect is relevant for comprehending weaknesses of such method when gathering data. In the case of DURATION, no difference between groups could be statistically identified, even though means and standard deviations give some support to the possibilities outlined above, where cluster 3 is associated with projects of shorter duration, while cluster 1 presents those projects with longer periods of development (cluster 2 stands in the middle). In the next section we will explore results for the 2006-2008, offering a robustness check for the validity of segments' structure.

Segmentation for 2006-2008

The analysis of clusters for the period 2006-2008 results in clusters of similar sizes in comparison to those found for projects finished within 2000-2005. Nonetheless, their internal structure differs significantly in terms of achieved results. We cannot exclude the chance of interference from different scales in the measurement of COMMACHIEV and EXP_IMPACT (see table 1 for description of variables).

The variable TECHACHIEV does not perform well in the process of cluster construction, as it can be seen by its lack of statistical significance (chi-square) for the three groups. The largest group in this case is cluster 3, which has worse outcomes in all of the levels under scrutiny. While its profile in technical terms is not properly bad, commercial achievements and expected impacts are fragile (and both are statistically significant regarding features of this cluster). This lack of marketability in face of relatively satisfactory technological leads us to classify this cluster under the Inventors category.

Cluster 2 is significantly defined only by its rate of commercial achievements. While TECHACHIEV suggests a very good profile in the technical dimension (similar to that of cluster 1 and above that of cluster 3), market results are poor, resulting in outcomes that are basically the same size of investments made. Expected impacts, though, are more optimistic, outperforming those observed in cluster 3, but not as positive as those in cluster 1. Such features, mainly the excellent capacity of achieving technological benefits, while "failing" in capturing economic benefits from them in the present should lead us to classify this cluster also as Inventors. Nonetheless, future expectations indicate the possibility of satisfactory results in the market. Hence, we define this cluster as Risky Innovators, even if its correspondence with Fischer and Molero's (2012) classification is imperfect.

Cluster 1 in 2006-2008 projects is the one which shows consistently better performances in the three dimensions used for cluster distribution (significantly represented by COMMACHIEV and EXP_IMPACT). We classify this cluster as Consistent Innovators. However, this classification does not represent perfectly the taxonomy proposed by Fischer and Molero (2012), since these companies outperform those from the other clusters in all of the aspects involved. An alternative would be to name such cluster as Successful Projects, where cluster 2 would be in an intermediate relative position, and cluster 3 would not represent failed projects necessarily, but those with the worst performance out of the three groups.

The analysis of differences between groups regarding a set of variables is presented below (table 6). As it can be noticed, none of such variables return significant results in terms of innovative intensity (RATIO_RD and TOT_COST), macroeconomic aspects (COUNTRY), and projects' dura-

tion (DURATION). We can then conclude that comparisons between datasets (2000-2005 and 2006-2008) do not perform well as a robustness evaluation. Besides differences in methods of data collection (Final Reports' structure), such finding indicates that results must be regarded carefully, as statistical findings are not consistent over time. While this is an indication of the weakness of the taxonomy under examination (Fischer and Molero, 2012), it captures the existence of relatively consistent behavioral patterns in firms, suggesting the validity of the methodological proposal contained in this research, even though its application in other contexts is uncertain before empirical assessments take place.

Concluding Remarks

Concerning the use of clustering techniques for the evaluation of firms' results according to relatively homogenous groups, this procedure is based on an analogous approach to that used for market segmentation. The underlying ra-

| Cluster Distribution | | Correspondence to Fischer and Molero's (2012) Classification | |
|----------------------------------|---|---|---|
| | Cluster 1 | 25 (21.7%) | Consistent Innovators |
| | Cluster 2 | 30 (26.1%) | Risky Innovators |
| | Cluster 3 | 60 (52.2%) | Inventors |
| | Missing | - | |
| Cluster Profiles and Main Trends | | | |
| | Cluster 1 | Cluster 2 | Cluster 3 |
| TECHACHIEV | Excellent (44%) and Good (52%) Technological Achievements | Excellent (40%) and Good (53.3%) Technological Achievements | Excellent (23.3%) and Good (58.3%) Technological Achievements. Also, 18.3% of Weak/Bad Attainments |
| COMMACHIEV | Commercial Achievements of 2-5 times RTD investment (76%) and +10 times RTD investment (20%)* | Commercial Achievements < or = RTD investment (100%)* | Very Low (100%) Commercial Achievements* |
| EXP_IMPACT | Expected Impacts of 2-5 times RTD investment (32%) and +10 times RTD investment (56%)* | Expected Impacts of 2-5 times RTD investment (56.7%) and +10 times RTD investment (26.7%) | Very low Expected Impacts (28.3%), < or = RTD investment (26.7%), and 2-5 times RTD investment (31.7%)* |

*Clusterwise Importance (chi-square at 95% confid.)

Percentages correspond to within cluster data.

Table 5. 2006-2008 Clusters

| <i>Means (Medians in parentheses; Std. Dev. in brackets)</i> | | | |
|--|-----------|-----------|-----------|
| | Cluster 1 | Cluster 2 | Cluster 3 |
| RATIO_RD | 2.16 | 2.28 | 2.19 |
| | (2) | (2) | (2) |
| | [.688] | [.702] | [.754] |
| COUNTRY | 2.12 | 1.96 | 2.23 |
| | (2) | (2) | (2) |
| | [.881] | [.889] | [.851] |

| <i>Non-Parametric Tests for Independent Samples</i> | | | |
|---|-----------|-----------|------|
| <i>Mann-Whitney Tests for Pairwise Comparison</i> | | | |
| | Cluster I | Cluster J | Sig. |
| RATIO_RD | 1 | 2 | .519 |
| | | 3 | .829 |
| | 2 | 1 | .519 |
| | | 3 | .628 |
| | 3 | 1 | .829 |
| | | 2 | .628 |
| COUNTRY | 1 | 2 | .521 |
| | | 3 | .580 |
| | 2 | 1 | .521 |
| | | 3 | .173 |
| | 3 | 1 | .580 |
| | | 2 | .173 |

| <i>Means (Std. Dev. in parentheses)</i> | | | |
|---|-----------|-----------|-----------|
| | Cluster 1 | Cluster 2 | Cluster 3 |
| TOT_COST | 2.47 | 2.84 | 2.23 |
| | (3.941) | (5.388) | (.851) |
| DURATION | 30.44 | 33.63 | 36.23 |
| | (15.798) | (15.294) | (12.795) |

| <i>One-Way Anova Results</i> | | | | |
|------------------------------------|-------------|-------------|---------------------|-------|
| <i>Bonferroni's Post Hoc Tests</i> | | | | |
| | Cluster (I) | Cluster (J) | Mean Difference I-J | Sig. |
| TOT_COST* | 1 | 2 | -.367 | 1.000 |
| | | 3 | -.086 | 1.000 |
| | 2 | 1 | .367 | 1.000 |
| | | 3 | .280 | 1.000 |
| | 3 | 1 | .086 | 1.000 |
| | | 2 | -.280 | 1.000 |
| DURATION** | 1 | 2 | -3.193 | 1.000 |
| | | 3 | -5.793 | .265 |
| | 2 | 1 | 3.193 | 1.000 |
| | | 3 | -2.600 | 1.000 |
| | 3 | 1 | 5.793 | .265 |
| | | 2 | 2.600 | 1.000 |

*Levene's sig. .954

**Levene's sig. .591

Table 36. 2006-2008 Clusters' ANOVA and Mann-Whitney tests

tionale is that policy intervention is more likely to succeed if it fits adequately groups with similar characteristics, instead of approaching heterogeneous groups of firms in a similar manner. In this research, this fragmentation of Eureka participants follows a post hoc structure, where the occurrence of relatively stable (and similar) groups of firms according to technological and market-related results seems to be more efficient than the use of firm-level information or macroeconomic conditions for formulation and adaptation of international R&D cooperation initiatives.

This particular methodology can function as a valuable instrument for policymakers in terms of international R&D networks monitoring and ex post intervention. The basic structure tested in this case is that of Fischer and Molero (2012) which divides companies in three different categories. Robustness of this framework was relevant for the 2000-2005 dataset, while it represented a fragile foundation for 2006-2008 projects. Nonetheless, present results represent a step forward in terms of clustering assessment of Eureka projects in comparison to that undertaken previously (Fischer and Molero, 2012), where a larger set of countries could be analyzed, more precise estimations could be offered and additional statistical tests were performed.

Each group presents strengths and weaknesses that can be supported through technical and managerial support. Risky Innovators lack steadiness in their ability to market results. Consistent Innovators represent companies with projects that could be addressed in terms of their incapacity to

exceed fair results (and achieve excellence). Inventors are those companies that could not meet their market goals, even in face of positive technological outcomes. As a control instrument, the cluster analyses facilitate evaluations, as well as further capacity to deal with such constraints in a combined manner (instead of a case-to-case approach). This tool also serves the purpose of identifying complementary agents that could be matched for future projects (according to their particular strengths).

Furthermore, this approach allowed a closer examination of influential variables in the determination of outcomes arising from Eureka's individual projects. We could verify differences between relatively homogeneous groups of firms regarding their inherent differences in a set of aspects. Quality of projects' management (FUNCTIONING) was significantly associated with higher rates of combined outcomes. This particular finding arises as an interesting collateral achievement of this research, highlighting the importance of managerial aspects related to transaction costs involved in (international) R&D cooperation. This particular finding represents an interesting tool for RTD policies that aim at fostering R&D collaboration, since it provides rather robust information that the managerial capacity of networks (which can be properly assessed ex ante via an assessment of the coordination structure of alliances) is strongly related to ultimate outcomes. Consequently, focus on projects that have a well-structured partnership will assure policymakers that value-for-money in these initiatives will be increased.

References

- AGHION, P.; David, P.A.; Foray, D. (2009). Science, technology and innovation for economic growth: linking policy research and practice in STIG systems. *Research Policy*, 38, 681-693.
- AGHION, P.; Tirole, J. (1994). The management of innovation. *The Quarterly Journal of Economics*, 109(4), 1185-1209.
- ARROW, K. (1962). Economic welfare and the allocation of resources for invention, in R. Nelson (ed) *The rate and direction of inventive activity*. Princeton University Press, Princeton, 609-625.
- AUDRETSCH, D.B.; Keilbach, M. (2008). Resolving the knowledge paradox: knowledge-spillover entrepreneurship and economic growth. *Research Policy*, 37, 1697-1705.
- BAYONA-SÁEZ, C.; García-Marco, T. (2010). Assessing the effectiveness of the Eureka Program. *Research Policy*, 39(10), 1375-1386.
- BENFRATELLO, L.; Sembenelli, A. (2002). Research joint ventures and firm level performance. *Research Policy*, 31, 493-507.
- BIN, A.; Salles-Filho, S. (2012). Science, technology and innovation management: contributions to a methodological framework. *Journal of Technology Management & Innovation*, 7(2), 73-86.
- BRANSTETTER, L.G.; Sakakibara, M. (2002). When do research consortia work well and why? Evidence from Japanese panel data. *The American Economic Review*, 92(1), 143-159.
- DURIEUX, L.; Fayl, G. (1997). The scheme used for evaluating the European research and technological development programmes. *Proceedings of the Conference on Policy Evaluation in Innovation and Technology*, June 25th to 27th, Paris, France, 373-381.
- EBERSBERGER, B.; Edler, J.; Lo, V. (2006). Improving Policy Understanding by Means of Secondary Analyses of Policy Evaluation: A concept development. *Fraunhofer ISI Discussion Papers – Innovation Systems and Policy Analysis*, Paper n. 12. Available at: http://www.isi.fraunhofer.de/isi-en/p/publikationen/diskpap_innosysteme_policyanalyse.php [Access in September, 2012]
- EUROPEAN COMMISSION (2002). *RTD Evaluation Toolbox: assessing the socio-economic impact of RTD policies*. Sevilla. Available at: http://ec.europa.eu/research/evaluations/pdf/archive/other_reports_studies_and_documents/assessing_the_socio_economic_impact_of_rtd_policies_2002.pdf [Access in May, 2012]
- EUROPEAN COMMISSION (2006). *Putting knowledge into practice: a broad-based innovation strategy for the EU*. Brussels.
- EUROPEAN COMMISSION (2006a). *Smart innovation: a practical guide to evaluating innovation programmes*, Brussels-Luxembourg. Available at: ftp://ftp.cordis.europa.eu/pub/innovation-policy/studies/sarl_smartinnovation_master2.pdf [Access in June, 2012]
- FISCHER, B.B. (2012). Methodological Lock-in and the evaluation of R&D policies: a critique to quasi-experimental assessments. *Current Opinion in Creativity, Innovation and Entrepreneurship*, 1(2).
- FISCHER, B. B.; Molero, J. (2012). Towards a taxonomy of firms engaged in international R&D networks: an evaluation of the Spanish participation in Eureka. *Journal of Technology Management & Innovation*, 7(3), 121-134.
- FREEMAN, C.; Soete, L. (2009). Developing science, technology and innovation indicators: what we can learn from the past. *Research Policy*, 38, 583-589.
- GARCÍA, J.V. (2010). A proposal of indicators and policy framework for innovation benchmark in Europe. *Journal of Technology Management & Innovation*, 5(2), 13-23.
- GEORGHIOU, L. (1999). Meta-evaluation: Evaluation of evaluations. *Scientometrics*, 45(3), 523-530.
- GEORGHIOU, L. (2001). Evolving frameworks for European collaboration in research and technology. *Research Policy*, 30, 891-903.
- GEORGHIOU, L. (2002). Impact and additionality of innovation policy. *Six Countries Programme on Innovation – Spring Conference 2002*, February 28th – March 1st. Available at: www.6cp.net/downloads/02brussels_georghiou_paper.doc [Access in July, 2012]
- GEORGHIOU, L.; Keenan, M. (2006). Evaluation of national foresight activities: assessing rationale, process and impact. *Technological Forecasting & Social Change*, 73, 761-777.
- GEORGHIOU, L.; Roessner, D. (2000). Evaluating technology programs: tools and methods. *Research Policy*, 29, 657-678.

- GILBERT, B.; Audretsch, D. B.; McDougall, P. P. (2004). The Emergence of Entrepreneurship Policy. *Small Business Economics*, 23, 313-323.
- KLETTE, J.; Moen, J. (1999). From growth theory to technology policy – coordination problems in theory and practice. *Nordic Journal of Political Economy*, 25, 53-74.
- KLINE, J.; Rosenberg, N. (1986). An Overview of Innovation. In: R. Landau and N. Rosenberg (eds). *The Positive Sum Strategies: harnessing Technology for Economic Growth*, National Academic Press, Washington D.C.
- KUHLMAN, S.; Edler, J. (2003). Scenarios of technology and innovation policies in Europe: investigating future governance. *Technological Forecasting and Social Change*, 70, 619-637.
- MALIK, K.; Cunningham, P. (2006). Transnational policy learning in Europe: attempts to transfer innovation policy practices. *Innovation: management, policy & practice*, 8, 262-272.
- MARÍN, P. L.; Siotis, G. (2008). Public policies towards Research Joint Venture: Institutional design and participants' characteristics. *Research Policy*, 37, 1057-1065.
- METCALFE, S. (1995). The design of order: notes on evolutionary principles and the dynamics of innovation. *Revue Économique*, 46(6), 1561-1583.
- METCALFE, S.; Georghiou, L. (1997). Equilibrium and Evolutionary Foundations of Technology Policy. CRIC Discussion Paper, Paper n. 3. Available at: <http://www.cric.ac.uk/cric/Pdfs/dp3.pdf> [Access in October, 2012]
- NELSON, R. R. (1959). The Simple Economics of Basic Scientific Research. *The Journal of Political Economy*, 67(3), 297-306.
- NELSON, R. R. (2007). Building Effective "Innovation Systems" Versus Dealing with "Market Failures" as Ways of Thinking about Technology Policy. Manchester Business School Working Paper, Paper n. 548. Available at: <http://www.mbs.ac.uk/research/workingpapers/index.aspx?AuthorId=4528> [Access in August, 2011]
- NELSON, R. R.; Winter, S. G. (2002). Evolutionary Theorizing in Economics. *Journal of Economic Perspectives*, 16(2), 23-46.
- NIST – National Institute of Standards and Technology – US (2003). A Toolkit for Evaluating Public R&D Investment: Models, methods, and findings from ATP's first decade. Gaithersburg, MD, US. Available at: <http://www.atp.nist.gov/eao/gcr03-857/contents.htm> [Access in February, 2011]
- PAVITT, K. (2003). The Process of Innovation. SPRU Electronic Working Paper Series, Paper n. 89. Available at: www.sussex.ac.uk/spru/documents/sewp89 [Access in June, 2012]
- RAYMOND, L.; St. Pierre, J. (2010). R&D as a determinant of innovation in manufacturing SMEs: an attempt at empirical clarification. *Technovation*, 30, 48-56.
- SANZ MENÉNDEZ, L. (1995). Policy choices, Institutional constraints and policy learning: the Spanish Science and Technology policy in the eighties. *International Journal of Technology Management*, 10, 622-641.
- SMITH, K. (2000). Innovation as a Systemic Phenomenon: rethinking the role of policy. *Enterprise & Innovation Management Studies*, 1(1), 73-102.
- SUURNA, M.; Kattel, R. (2010). Europeanization of innovation policy in Central and Eastern Europe. *Science and Public Policy*, 37(9), 646-664.
- US DEPARTMENT OF ENERGY (2007). Overview of evaluation methods for R&D programs: a directory of evaluation methods relevant to technology development programs. Albuquerque. Available at: http://www1.eere.energy.gov/ba/pba/pdfs/evaluation_methods_r_and_d.pdf [Access in January, 2012]