Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia - Social and Behavioral Sciences 75 (2013) 353 – 358

Procedia
Social and Behavioral Sciences

2nd International Conference on Leadership, Technology and Innovation Management

Knowledge search strategies in the biotechnology domain: a patent benchmarking analysis

Corrado lo Storto*

University of Naples Federico II, Piazzale V. Tecchio 80, Naples 80125, Italy

Abstract

This paper offers a new interpretive perspective to investigate factors of researchers productivity, moving from a personal features/resource endowment view to a behavioral one centered on the knowledge search. Little work has indeed been carried out on the relationship between the research productivity and the scientist knowledge search behaviors. Three knowledge search dimensions are taken in account: search type, search focus and search dynamics. Using data relative to 873 biotechnology patents granted from 1960 to 2007 to 255 academic scholars that are affiliated to 36 Italian universities, this paper investigates if a particular knowledge search behavior is associated to greater patenting rates.

Keywords: Knowledge, Search, Biotechnology, Patents, University, Scientists, Research productivity, Benchmarking

1. Introduction

Since the 1970s, the university research has played a particular important role both for the advancement and dissemination of the stock of scientific knowledge and the growing of industry in the biotechnology field. Indeed, biotechnology is a science-based industry and an extraordinary example of effective interaction between basic and applied research (Audretsch and Feldman, 1996; McMillan et al., 2000; Cockburn, 2005). It is not infrequent that new companies are formed as academic spin-offs (Audretsch, 2001; Chiesa and Piccaluga, 2000; Walsh et al., 1995). Traditionally, the stock of knowledge has been freely available within the research community, accessible through journal articles, conference papers, databases, scientists mobility, etc. But, more recently with the increasing awareness of the commercial potential of the biotechnology research outlets and the need for funding of the academic institutions, university scientists have been more and more involved in the exploitation of the research activity through patenting. Although Italian biotech is characterized by recognized scientific excellence and a high degree of dynamism, Italy is ranked fourth as to the overall number of scientific publications in the biotechnology domain (close to 6%) preceded by US, UK, and Germany, but it ranks only fifth, below France, in terms of patents productivity, with a relative share of 3% of the total number of patents granted in Europe or 2% of patents granted in US (Pammolli, 2011).

In this paper, I will analyze patterns of knowledge search in Italian universities active in the biotechnology field. In particular, by using data relative to 873 biotechnology patents granted from 1960 to 2007 to 255 academic scholars

* Tel.: +39-081-768.2932 fax: +39-081-768.2154

E-mail address: corrado.lostorto@unina.it

doi:10.1016/j.sbspro.2013.04.040

affiliated to 36 Italian universities, this paper investigates if there is an association between the knowledge search behavior pursued by a researcher and his/her patenting productivity. Three knowledge search dimensions are considered to model the researchers knowledge search behavior: search type, search focus and search dynamics. Indexes of technological familiarity, technological concentration, and collaboration are built and measured to qualify search behaviors.

2. Modeling scientific research as knowledge search

As a form of problem-solving activity, scientific research involves the generation and recombination of ideas, concepts, insights, interpretation of data and information that lead to the creation of new knowledge (Katila and Ahuja, 2002). Indeed, knowledge does not have a rigid nature, but it can be transformed, combined, accumulated, stored and transferred. The search for new knowledge useful to solve old unsolved problems or identify new problems to deal with represents the principal activity of scientific research. The search for new knowledge is mostly an evolutionary path-dependent, intrinsically uncertain and interactive process, and scientists – due to the resource and cognitive limitations that make it difficult to formulate objective prior beliefs - tend to pursue their innovative search for new knowledge and solutions to problems leveraging on the knowledge stock they are more familiar with or within a set of knowledge domains that are very close to the familiar knowledge domain which are many times freely available in the academic and technical communities, refining and improving it, or, as frequently occurs, combining them together (March, 1991). When scientists become engaged in problem-solving activity in search of new useful knowledge, they use existing knowledge as a recipe, using as ingredients the elementary components of knowledge information, physical components and processes (Nelson and Winter, 1982). This recipe shows how these ingredients should be combined to obtain useful outlets. Scientific research is thus a process in which scientists search for new recipes, i.e. new ways useful for the recombination of single pieces of knowledge to generate new and better alternatives. When researchers accumulate expertise in a specific knowledge domain, they prefer rely upon this expertise even when they deal with new problems that require a change in the knowledge search perspective, thus limiting the potential set of alternatives, setting more or less rigid constraints to the search activity. When scientists either recombine the components of a familiar set of knowledge pieces or refine single or combinations of knowledge pieces previously used, they move within a bounded search space, reducing uncertainty and limiting risk, carrying on local search that becomes a synonymous of exploitation. On the contrary, when scientists try either new knowledge combinations or new single knowledge components, they carry on distant search or exploration (March, 1991; Simon, 1991). The spreading out of the range of knowledge search help researchers finding useful solutions distant from the commonly exploited reference models, discovering alternatives that were disregarded in the past (Brabazon and Matthews, 2002).

3. The study setting

Data relative to patents granted between 1960 and 2007 to 255 Italian university researchers actively working in the biotechnology field were collected for the empirical study from the EPO databank. In total, 873 granted patents were considered for the analysis, while the single university researcher was the primary research unit. The following procedure was implemented to assemble the patents sample. Firstly, patents related to biotechnology were classified developing an a priori list of biotechnology related fields of interest based on the OECD classification (van Beuzekom and Arundel 2006), associating a number of IPC classes and subclasses to each field. Secondly, researchers working in the biotechnology field in Italian universities were identified and adopting an approach similar to the one proposed by Trajtenberg (2004), the EPO databank was queried to check whether the specific researcher had been granted at least one patent classified as a biotechnology patent. Researchers include: 43 assistant professors, 69 associate professors, and 143 full professors. Finally, data relative to patents granted to every inventor in the biotechnology field were collected and a databank that associates 255 inventors names to 873 biotechnology patents distributed on a temporal window from 1960 to 2007 was constructed. Patents granted to more than one researcher were not included in the sample. All the assignees of patents in the sample are individual researchers. Indeed, according to the Italian law, the university researchers do not have to comply with the duty of the employer's ownership of inventions.

A forward stepwise statistical regression analysis was performed to measure how specific knowledge search behaviors affect scientists research productivity (Stevens, 1996). This method identifies the independent variables that account for the highest proportion of the observed variance of the dependent variable, and – as many scholars point out - it is often the best compromise to find an equation capable to predict the maximum variance for a specific dataset (Hocking, 1976).

3.1. Research productivity as dependent variable

The count of patents granted from 1960 to 2007 to every researcher included in the sample was used to measure research productivity. A semi logarithmic transformation of this variable was used as dependent variable, assuming an exponential trend (LOGPAT).

3.2. Variables modeling the scientists knowledge search behavior

According to the IPC international patent classification standard, a patent can be identified as the combination of a number of technological components associated to certain classes and subclasses. These latter can be indirectly used to investigate the type of knowledge search pursued by the inventor. Moreover, as the patent databank contains information relative to time when each patent was granted, moving back over time, the knowledge search evolution can be reconstructed, and the innovation search trajectories can henceforth be explored.

Three dimensions of the knowledge search behaviour were considered in this study: 1) the *search scope*, which classifies search as being either exploitative or explorative; 2) the *search focus* that classifies search in terms of its being more or less dispersed across different knowledge domains; and 3) the *search dynamics*, that views search in terms of its being more or less uniformly distributed over time. Consistently, three types of patent-based indexes were used to investigate the knowledge search behavior of every firm: two familiarity indexes (FS and FC), the technological concentration index (TC), and two inequality indexes (GFSI and GFCI).

Search scope. The FS^i_j index measures the familiarity an inventor has with the technological component j, the subclass j of IPC standard of a patent i (Fleming, 2001; lo Storto, 2006). The greater its value, the greater the familiarity with component j. FS^i_j is calculated as a product between two factors, the first one obtained as a summation of the number of all patents which were codified with that particular subclass before the reference year, and the second one used as a damping factor that takes into account for the loss of knowledge over time:

$$FS^i_j = \sum_{\substack{\text{all patents k granted} \\ \text{before patent i}}} 1\{\text{if patent k uses subclass } j\} \times e^{-\left(\frac{ti-tk}{c}\right)}$$

t_i = year in which patent i was granted
t_k=year in which patent k was granted
c= knowledge loss constant

FS^i_j was calculated for every patent subclass. Therefore, a patent classified into n subclasses will have n values for the index. Finally, for every year a FS was calculated as an averaged summation of all FS^i_j indexes (AFS). The rationale behind the calculation of FS is that the use of a familiar component will enable an inventor to build new knowledge leveraging on the efforts made to accumulate useful knowledge available in the scientific and technological field, avoiding to use not effective components.

The FC^i_j is a measure of the inventor familiarity for a combination j of technological components included in a patent i (Fleming, 2001; lo Storto, 2006).

$$FC^i_j = \sum_{\substack{\text{all patents granted} \\ \text{before patent i}}} 1\{\text{if patent k uses a combination of subclasses } j\} \times e^{-\left(\frac{ti-tk}{c}\right)}$$

t_i = year in which patent i was granted
t_k=year in which patent k was granted
c= knowledge loss constant

The index has a structure similar to that of the FS index. The greater FC^i_j, the greater the familiarity with combination j. A higher value of this index also indicates a progressive refining of the combination used in a patent. FC^i_j increases when the frequency in the usage of a specific combination of components increases and its use is almost recent. The FC^i_j index was calculated for every combination of a patent subclasses. A global FC was calculated as an averaged summation of all FC^i_j indexes in every year (AFC).

For both FS and FC indexes, the value of 5 was set for the constant c (Fleming, 2001; lo Storto, 2006).

Search focus. The TC index measures to what extent the innovation search of the inventor is focused on a single technological subcomponent - the IPC subclass. This index ranges between 0 and 1, this latter value achieved if the inventor is granted patents in only one technological subclass (lo Storto, 2006):

$$TC_{subclass} = \sum_j \left(\frac{N_j}{N} \right)^2$$

N = number of patents granted to an inventor
j = patent subclass

Search dynamics. The GFSI and GFCI respectively measure the inequality of the FS and FC distributions over time in the research activity dynamics of the firm. These indexes have the same structure, based on the GINI index (Gini, 1936). As to FS, it is:

$$GFSI = 1 - \sum_{i=i_1}^{48} \left(FS_i + FS_{i-1} \right)\left(y_i - y_{i-1} \right)$$

FS_i = the cumulated proportion of the FS variable, for $i = i_1,...,48$
y_i = the cumulated proportion of the time variable, for $i = i_1,...,48$
48 = the number of years included in the temporal window
i_1=the year in which the first patent was granted to the inventor

A low value of the index is associated to a more uniform distribution over time of FS or FC; vice versa, a higher value indicates a more unequal time distribution of knowledge search.

3.3. Additional (control) variables

Four additional variables have been added to the model:

Department size (DEPTSIZE). The university department size was measured by the total number of researchers employed in the researcher j department, including both those scientists working in the biotechnology field and those not working on that field. The inclusion of this variable is justified by empirical evidence suggesting that research productivity increases with team size.

Researcher academic position (POSITION). A dummy 3-level variable was included in the regression analysis to take into account for the researcher academic position. Variable code are: 1 for assistant professors, 2 for associate professors and 3 for full professors. The addition of this variables has ist justification in empirical evidence that found that academic rank might be a predictor of research productivity (Fulton and Trow, 1974).

Collaboration between inventors (IC). Empirical literature supports the idea that collaborative research between academic researchers and other academic researchers and between academic researchers and industry researchers increases the academic researchers productivity allowing cross fertilization of ideas (Siegel et al, 2003; Powell et al., 2005). An index (IC_j) measuring the inventor j collaboration intensity with other researchers inside or outside the affiliated department during the innovation activity was included in the regression model. The higher the index value, the higher the collaboration intensity.

$$IC_j = \frac{\sum_{i=1}^{m} number\ of\ co\text{-}inventors\ for\ patent\ i}{m}$$

m = number of patents granted to inventor j

Inventing tenure (AGE). A variable that measures the inventing tenure of every scientist was also included in the statistical analysis. This variable was built as the number of years elapsed from the final end of the temporal window of analysis (2007) to the year when the scientist was granted his/her first patent.

4. Results

Table 1 shows the outcome of the stepwise regression analysis. Variables values have been preliminarily normalized by dividing them by the maximum. Figures reveal that not all variables included in the regression model affect research productivity of scientists. In particular, it seems that neither the academic rank of the researcher nor the collaboration intensity affect his/her patenting productivity. Contrarily to what expected and reported in literature, an increase in the department size negatively influence the number of patents granted to researchers, even though this impact is modest. As expected the effect of the inventing tenure on the researcher productivity is positive, but not relevant. As to variables describing the scientist knowledge search behavior, only for the search focus (TC) the effect on patenting productivity is clearly identified. This effect is negative and remarkable, and an increase in the focus of the research effort on one (one single IPC subclass) or a limited domain of knowledge components (a small number of IPC subclasses) pushes down the researcher productivity. As to the knowledge search scope, the average familiarity for the technological subclass (AFS) has no significant statistical influence on the patenting productivity. On the contrary, the average familiarity for the combination of subclasses (AFC) positively affects the number of patents granted to researchers. However, this effect is not relevant. The analysis of findings relative to the knowledge search dynamics are also intriguing. Indeed, there is no effect on the patenting productivity of the FC index temporal distribution, but the dynamics of the FS over time largely influence the researchers productivity in terms of patents granted. This effect is negative, as an unequal distribution of the familiarity index relative to the single technological subclasses over the inventing life of the scientist decreases the patenting productivity.

Table 1. Stepwise regression analysis. LOGPAT as dependent variable

effect		level of effect	comment	param.	beta	t	prob.
intercept				16.377		25.045	0.000
DEPTSIZE				-0.239	-0.045	-2.375	0.018
POSITION		1	pooled				
POSITION		2	pooled				
IC			pooled				
AGE				0.449	0.085	4.159	0.000
TC				-1.079	-0.451	-20.187	0.000
AFS			pooled				
AFC				0.284	0.052	2.356	0.019
GFSI				-15.214	-0.588	-22.845	0.000
GFCI			pooled				
adjusted r-squared	0.914						
F	517.722						
prob.	0.000						

5. Conclusion

This paper has presented a framework that explains a scientist patenting productivity in terms of his/her knowledge search behavior. Three dimensions of the knowledge search behavior are taken into account: search type, search focus and search dynamics. In particular, findings show that:

- the knowedge search scope has an impact on the patent productivity rate. A strong focalization of the knowledge search behavior on a limited knowledge domain may be detrimental to research productivity.
- the strong negative and statistically significant value of GSFI emphasizes that how knowledge search is distributed along the academic life of a researcher together with the knowledge search scope (exploitative vs explorative) may affect research productivity. Particularly, a discontinuous, very unbalanced and not uniformly distribution of search relative to single knowledge components (IPC subclasses) has a negative effect upon research productivity. Researchers who adopt a continuous and more uniform search behavior balancing distant and local search are more productive in terms of patents granted. That is consistent with the insights from the analysis relative to AFC. Researchers that use recombination of familiar technology components to develop new knowledge have a higher patenting productivity. Thus, an incremental exploitative approach to knowledge development seems more conducive to higher patenting productivity.
- collaboration is not critical for the patenting productivity. This outcome remains equivocal. A more in depth investigation of how collaboration between researchers influences research productivity is needed, as literature emphasize how this variable has a positive effect on research productivity (Jones and Preusz, 1993). Scientific production is not so much the result of a single researcher's effort, but rather the outcome of a cognitive process

which involves a number of researchers and the whole scientific community. In particular, interaction limited only to university researchers should be distinguished from interaction between university and industry researchers. As Siegel et al. (2003) show, interaction with industry can often lead to the development of new research ideas and concepts, and consequently enhanced productivity. Interaction of academic researchers with industry can lead to an increased interest in patenting, both because industry is more interested in technology transfer and commercial applications of research output (e.g., patents), and because it directs its research to questions that are well suited to patenting (Agrawal and Henderson, 2002). How scientists interact with one another should also be taken into account as the quality of interaction affects the amount of social exchange due to social interactions and, hence, their scientific productivity. In addition, team homogeneity (or heterogeneity) is an important factor to be considered. A high researchers homogeneity may have perverse effects on creativity and research productivity.

References

Agrawal, A. and Henderson, R. (2002), Putting patents in context: Exploring knowledge transfer from MIT, Management Science, 48, pp. 44-60.

Audretsch, D.B. and Feldman, M.P. (1996), R&D spillovers and the geography of innovation and production, American Economic Review, 86, pp. 253-273.

Audretsch, D.B. (2001), The role of small firms in U.S. biotechnology clusters, Small Business Economics, 17, pp. 3-15.

Brabazon, T. and Matthews, R. (2002), Technological invention as recombinant search, Working Paper, JEL classification 029-031, University College Dublin, pp. 1-22.

Chiesa, V. and Piccaluga, A. (2000), Exploitation and diffusion of public research: the case of academic spin-off companies in Italy, R&D Management, 30, pp. 329–340.

Cockburn, I. (2005), State street meets the Human Genome Project, in Hahn, R. (ed.) Intellectual Property Rights in Frontier Industries: Bioinformatics and Software, Washington: Brookings-AEI Joint Center.

Fleming, L. (2001), Recombinant uncertainty in technological search, Management Science, 47, pp. 117-132.

Fontes, M. (2001), Biotechnology entrepreneurs and technology transfer in an intermediate economy, Technological Forecasting and Social Change, 66, pp. 59-74.

Fulton, O. and Trow, M. (1974), Research activities in American higher education, Sociology of Education, 47 (Winter), pp. 29-73.

Gini, C. (1936), On the measure of concentration with special reference to income and statistics, General Series, No. 208, Colorado College Publication, pp. 73-79.

Hocking, R.R. (1976), The analysis and selection of variables in linear regression, Biometrics, 32, pp. 1-49.

Jones, J.E. and Preusz, G.C. (1993), Attitudinal factors associated with individual faculty research productivity in academic dentistry, Perceptual and Motor Skills, 76, pp. 1191-1198.

Katila, R. and Ahuja G. (2002), Something old, something new: A longitudinal study of search behavior and new product introduction, Academy of Management Journal, 45, pp. 1183-1194.

lo Storto, C. (2006), A method based on patent analysis for the investigation of technological innovation strategies: the European medical prostheses industry, Technovation, 26, pp. 932-942.

March, J. (1991), Exploration and exploitation in organizational learning, Organization Science, 2, pp. 71-87.

McMillan, G.S., Narin, F. and Deeds, D.L. (2000), An analysis of the critical role of public science in innovation: the case of biotechnology, Research Policy, 29, pp. 1-8.

Nelson, R. and Winter, S. (1982), An evolutionary theory of economic change, Cambridge, MA: Belknap Press.

Pammolli, F. (2011), Il valore economico delle scienze della vita, Research Report, IMT Lucca.

Powell, W.W., White, D.R., Koput, K.W. and Owen-Smith, J. (2005), Network dynamics and field evolution: The growth of interorganizational collaboration in the life Sciences, American Journal of Sociology, 110, pp. 1132-1205.

Siegel, D.S., Waldman, D. and Link, A. (2003), Assessing the impact of organizational practices on the relative productivity of university technology transfer offices: An exploratory study, Research Policy, 32, pp. 27-48.

Simon, H.A. (1991), Bounded rationality and organizational learning, Organization Science, 2, pp. 125-134.

Stevens, J. (1996), Applied multivariate statistics for the social sciences (3rd ed.), Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Trajtenberg, M. (2004), The names game: Using inventors patent data in economic research, Tel Aviv University, NBER and CEPR Working Paper, June.

van Beuzekom, B. and Arundel A. (2006), OECD Biotechnology statistics – 2006, Paris: OECD.

Walsh, V., Niosi, J. and Mustar, P. (1995), Small-firm formation in biotechnology: A comparison of France, Britain and Canada, Technovation, 15, pp. 303-327.