STP

SCIENCE AND TECHNOLOGY POLICIES RESEARCH CENTER

Working Paper Series

07/01

Breadth and Depth of Main Technology Fields: An Empirical Investigation Using Patent Data

Müge Özman
Science and Technology Policies Research Centre
Middle East Technical University

Science and Technology Policies Research Center
Middle East Technical University
Ankara 06531 Turkey
http://www.stps.metu.edu.tr

Breadth and Depth of Main Technology Fields: An empirical investigation using patent data

Muge Ozman
Middle East Technical University
Science and Technology Policy Studies Research Centre
e-mail: ozman@cournot.u-strasbg.fr

June 6, 2007

Abstract

Recently work on technological change has emphasized the importance and implications of different knowledge bases among industries in terms of innovative potential. In some industries, products and processes have become more complex, as well as there appears to be increased convergence in some market segments. Although increasing attention has been given to features of knowledge bases, there have been limited empirical research on how to measure them. In this paper a method is proposed to measure empirically the breadth and depth dimensions of main technology fields, sectors and firms in the economy. For this purpose, we measure the knowledge bases of 30 main technology fields by using the concepts of breadth and depth. Breadth corresponds to the range of different subjects that a technology field draws upon. Depth refers to the extent to which a certain field is exploited in detail. We position the main technology fields in the breadth and depth space by utilizing the EPO (European Patent Office) database between 1978-2000. We also present the evolution of breadth and depth through time, as well as the breadth and depth dimensions of 40 largest firms in biotechnology and telecommunications. Our results reveal that the both technology fields and firms are largely scattered in the breadth and depth space. Biotechnology stands out to have the highest breadth and depth.

Key Words: Patents, Breadth and depth, technology fields, knowledge base.

1 Introduction

Recently, work on technological change has emphasized the importance and implications of different nature of knowledge bases of both industries and firms in terms of their effect on innovative performance (Nesta and Saviotti, 2004) and organizational structure (Orsenigo et al., 2000; Dosi and Hobday, 1999; Prencipe, 2000; Brusoni and Prencipe, 2000; Ozman, 2005; Cowan et al., 2004). Research in this area suggests that in some industries, products and processes have become more complex, as well as there appears to be increased convergence in some market segments, which obviously influences the efficient choice of organizational structure. For example the literature on complex product systems stresses the importance of the harmony between complexity in products and organization structure. In addition, an expanding area of research is concerned with the characteristics of the knowledge base of firms as well as industries, drawing a distinction between technology and product boundaries (Brusoni and Prencipe; 2000; Brusoni et al., 2001). Nesta and Saviotti (2004) analyse the diversity and coherence of knowledge base of pharmaceutical firms and conclude that both has a positive and significance influence on innovative performance of the firm.

In the literature, complexity in the knowledge base have been taken in various ways. One of the commonly used frameworks is related to the number of components and interdependence among them (Simon, 1969; Zander and Kogut, 1995; Kauffman, 1993) Among these, Wang and Tunzelmann (2000) define two dimensions of complexity as breadth and depth. Complexity in depth refers to "analytical sophistication of a subject which becomes complex because of the cognitive difficulty in pushing the particular matter to its logical extremes" (p.806). Whereas complexity in breadth refers to "the range of areas that have been investigated to develop a particular subject". In short depth is concerned with the level of sophistication and breadth is concerned with the level of heterogeneity. Defined in this way, product complexity in breadth refers to the embodiment of a larger number of components/assemblies which makes up a product, whereas depth dimension refers to the degree of cognitive complexity embodied in these components. Therefore changes in complexity in one dimension does not necessarily influence complexity in the other.

Although increasing attention has been given to the characteristics of knowledge bases in recent years, there have been relatively limited attempts to measure features of the knowledge bases empirically. This paper is an attempt in this regard. It is believed that empirical research in this area is very limited and highly valuable, to permit studying the effect of knowledge bases on innovative performance and also organizational structures.

The aim of the paper is to present a methodology to measure the breadth and depth dimensions of main technology fields, and apply this methodology to 30 main technology fields. The same methodology is used to measure the breadth and depth dimensions of 40 largest firms in biotechnology and telecommunications sectors. In the context of the study, the following questions are addressed: How have the knowledge bases of the main technology fields evolved over time? How do firms compare to each other in terms of the breadth and depth of their knowledge bases?

To address the above questions, the patent document is taken to be a product of innovation efforts, and the main technology fields are positioned in the breadth and depth space by utilizing the EPO database between 1978-2000. A measure of breadth and depth is constructed for each patent taken in the mentioned period according to technology field. The first part of the paper is concerned with the description of how breadth and depth measures are derived. Breadth is measured by the range of different technology fields that a patent includes in its International Patent Classification (IPC) list. The depth of the patent is measured by the extent to which a patent draws upon a certain field more intensively than others, also measured by using the IPC list of the patent. the evolution of breadth and depth in some sectors and the firm based breadth and depth values are provided. In the second part of the paper the breadth and depth of main technology fields, their evolution through time, and the breadth and depth dimensions of 40 largest firms in biotechnology and telecommunications are presented.

2 Breadth and Depth of Main Technology Fields

In this section, the analysis of breadth and depth in 30 main technology fields, the evolution of breadth and depth in selected knowledge intensive technology fields, as well as an analysis of the breadth and depth dimensions of largest firms in two sectors, namely biotechnology and telecommunications are presented.

2.1 Data

The EPO/CESPRI database includes all the patents granted by the EPO between the years 1978 and 2000. A total of 668,947 patents were included, excluding those that do not have a secondary IPC code. A patent document is a rich source of information, since it contains information on the relevant technology codes related with the subject matter of the patent, which is given by the 8-digit IPC code. A patent document is assigned a main IPC code, as well as secondary codes. The patents in the sample are partitioned into 30 main technology fields, based on their main IPC codes. This classification has been prepared by Fraunhofer Gessellschaft-ISI (Karlsrube), Institut National de la Propriete Industrielle (INPI-Paris) and Observatoire des Sciences et des Techniques (OST, Paris) and it is mainly composed of an allocation of IPC codes into technology fields. The distribution of patents in the sample period, according to technology fields are given in Figure 1.

Before proceeding with the explanation of the breadth and depth indices, it is useful to briefly mention the structure of the IPC system.

2.2 Structure of International Patent Classification (IPC)

In the IPC system, there are 8 sections revealed by the first digit of the code; classes, which are revealed by the first 3 digits, which are in turn divided into subclasses; and groups, as revealed by the first 6 digits. An example is A21B01/06. In this example, A is the section which is covered under the heading of human necessities. A21 is the class (baking) and A21B is the subclass (bakers ovens, machines or equipment for baking). A21B01 is the group level and A21B01/06 is the subgroup level, which

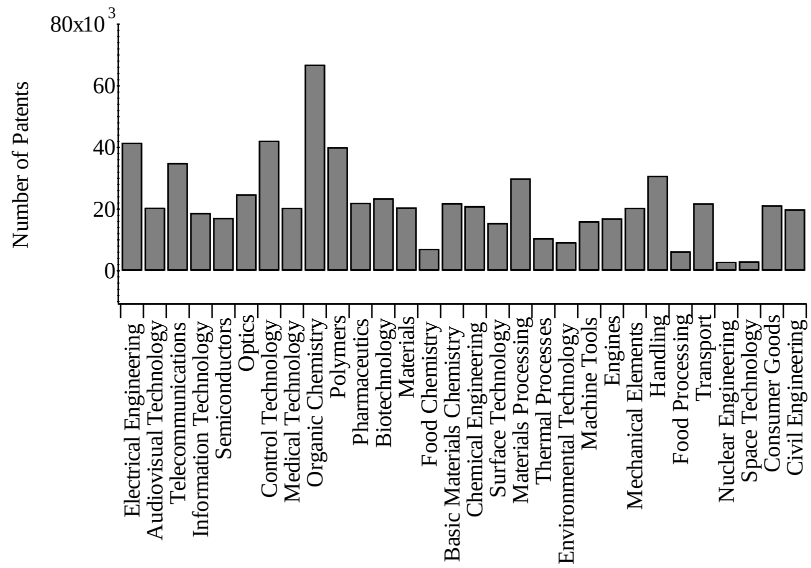

Figure 1: Number of Patents in the Database

corresponds to bakers ovens heated by radiators. Each patent granted by the EPO has a main IPC code, and secondary codes. To calculate the breadth and depth measures,the IPC list of the patent in both main and secondary codes are utilized. The main IPC code of the patent is used to assign the patent to one of the technology fields, and the secondary codes are used to measure the breadth and depth dimensions of the patent in terms of its coverage.

2.3 Breadth

The breadth of the patent measures the width of different technology fields that the secondary IPC list of a patent in a certain technology field covers. The more the different types of technology fields a given patent extends to, the wider its knowledge base taken to be. For this purpose, several different measures are used, to test for robustness of the results.

The first measure is based on a direct count of the number of secondary IPC codes of a patent in different technology fields, other than its own technology field (patents own technology field is given by its main IPC code). Formally, let $I = \{1.....30\}$ represent the 30 technology fields, and $\Gamma_{ij} = \{k \in I \ \setminus \ \{j\} \mid x_i(k) = 1\}$ represent the set of technology fields of the IPC list of the patent i which is in technology field j, and $x_i(k) = 1$ if the patent has an IPC code in field k. Then the breadth of this patent is given by $b_{ij} = \# \ \Gamma_{ij}$. The breadth of the technology field is the average of the breadth over all patents in this technology field;

$$B_j = \frac{\sum b_{ij}}{N_j}$$

where N_j is the total number of patents in technology field j.

One disadvantage of this measure is that it assigns an equal weight to all the technology fields with respect to their relatedness. In other words in devising a breadth measure, one should take into account the relative relatedness between the technology field of the patent and the other technology fields that it covers. For example, the technology fields biotechnology and pharmaceuticals are relatively more related to each other than pharmaceuticals and electrical engineering. In this case, if in the IPC code of a pharmaceuticals patent there exists an IPC code that belongs to electrical engineering, this field should have a higher weight in the breadth, than an IPC code which belongs to biotechnology. In the second measure that was developed, the relatedness among technology fields were taken into account, based on the relatedness measures devised by Breschi et al. (2003). Here, the relatedness among technology fields is calculated based on the citation rates.

For a patent i in technology field j, weighted breadth is:

$$b_{ij} = \sum_{k \in I} x_i(k)(1 - R_{jk})$$

where R_{jk} is the relatedness between fields $j$and k.[1]

Figure 2 reveals that, the technology fields are significantly different from each other in terms of their breadth measures. In most fields, the breadth measure increases when it is weighted with respect to relatedness, with the exception of biotech-

[1]See Breschi et. al (2003) for the matrix of relatedness weights among 30 technology fields.

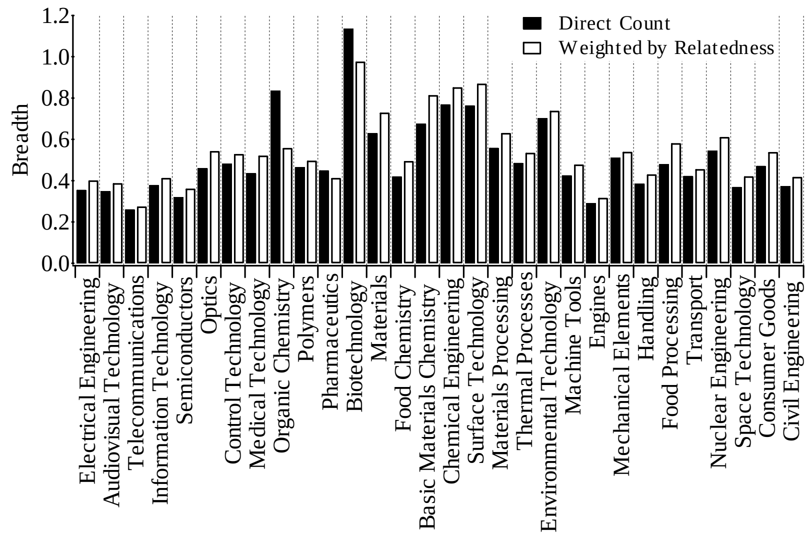

Figure 2: Weighted and Direct Count Breadth Measures for 30 Main Technology Fields

nology and organic chemistry where the breadth measures decrease, mainly due to the fact that although the patents in these fields have a significant average number of IPC codes in other technology fields, these are mainly in highly related areas. Overall, it can be seen that the relative breadth of fields remain constant even when adjusted for relatedness, which basically confirms the robustness of our breadth measures. In the rest of the paper the weighted measures are used.

2.4 Depth

The depth measure is conceptually less straightforward than breadth. The depth of a patent refers to the extent to which a patent is specialized in a certain field. Therefore, a patent may have both high breadth and high depth, if, for example, it draws upon a wide range of different aggregated fields (which will be reflected in the breadth measure), but also in a certain aggregate field it utilizes a high number of detailed fields.

The first measure of depth that is employed is based on the number of IPC codes of the patent which are the same as the technology field of the patent itself. Formally, let $I = \{1.....30\}$ represent the 30 technology fields. If patent i is in technology field j, $\Gamma_{ij} = \{k^j \in I \ \mid x_i(k^j) = 1\}$ represent the set of IPC codes of the patent which is in technology field j, ($x_i(k^j) = 1$ if the patent has a secondary IPC code in field j), namely its own technology field as revealed by the main IPC code. Then the depth of this patent is given by $d_{ij} = \# \, \Gamma_{ij}$. The depth of the technology field j is the average of the breadth over all patents;

$$D_j = \frac{\sum d_{ij}}{N_j}$$

where N_j is the total number of patents in technology field j. Therefore, the depth measures the extent to which the patent is concerned with its own technology field. Although this measure provides a rough idea on the analytical sophistication of the patent, a more precise measure should take into account the structure of the IPC codes themselves. In other words, the IPC list of the patent may include many IPC codes from the same technology field, but would this mean that the patent is "deep"? Within its own technology field, the patent may still be drawing upon different knowledge types rather than exploiting a certain knowledge unit in depth. To capture this, a more detailed analysis of the specific IPC codes of the patent was made.

An important point here is that two IPC codes can differ in many levels, including section level (e.g. **A**01N57/12 and **B**03B01/10), class level (**A01**N57/12 and **A02**L01/00), subclass level (**A01N**57/12 and **A01L**01/10), group level (e.g. A01N**57**/12 and A01N**47**/10), subgroup level (A01N57/**12** and A01N57/**09**). This hierarchical categorization gives an idea about the extent to which the subject matter of the patent is exploiting a certain field in depth. In particular, a high number of IPC codes in lower levels of hierarchy implies that the more refined the subject matter of the patent is. Let us consider two extreme cases. In high levels of hierarchy, **A**01N57/12 and **B**03B01/10 differ in the section level, if the patent includes these two codes at the same time in its IPC list, it means that it draws upon two very different knowledge areas. On the other hand, in the IPC list of the patent, if we see A01N57/**09** and A01N57/**12** at the same time, the patent draws upon one specific knowledge type in detail, with many aspects of it, namely the area A01N57. According to this approach, the repetition of lower hierarchy levels reveal increasing depth of the patent.

Taking this into account, the above depth indices were weighted, with respect to an index of dominance that was calculated on the 6 digit level. As mentioned above, the depth of the knowledge base is defined to be the dominance of some knowledge types in the patent. For this purpose, a depth index is constructed that measures the extent to which the patent is specialized, e.g., relying intensively on a small number of subjects.

To construct an index of depth, the frequency that each subgroup repeats in the IPC list of the patent, and thus the percentage of this subgroup was calculated. It is assumed that the higher is the weight of particular subgroup, the deeper is the patent.

To calculate a measure of dominance of a certain subgroup, the Blau index (1977) was utilized. For a patent i in technology field j, the index is calculated as follows:

$$w_{ij} = 1 - \sum a_{ik}^2$$

where a_{ik} is the proportion of a certain subgroup k in patent i. Higher values of the Blau index indicate that all knowledge types are utilized in similar proportions, whereas smaller values indicate the intensive use of some knowledge types over others. Therefore, *lower* values indicate a *deeper* knowledge base (more specialized). One disadvantage associated with this index is that depending on the breadth of the patent, the maximum and minimum of the index is different, so that the index cannot be used as it is to compare patents of different breadth measures. This is to say that, when there are a higher number of different subclasses, then this will be reflected in a larger possible maximum value of the index. To be able to make the depth measures comparable, the depth measures obtained is divided by the number of items in the

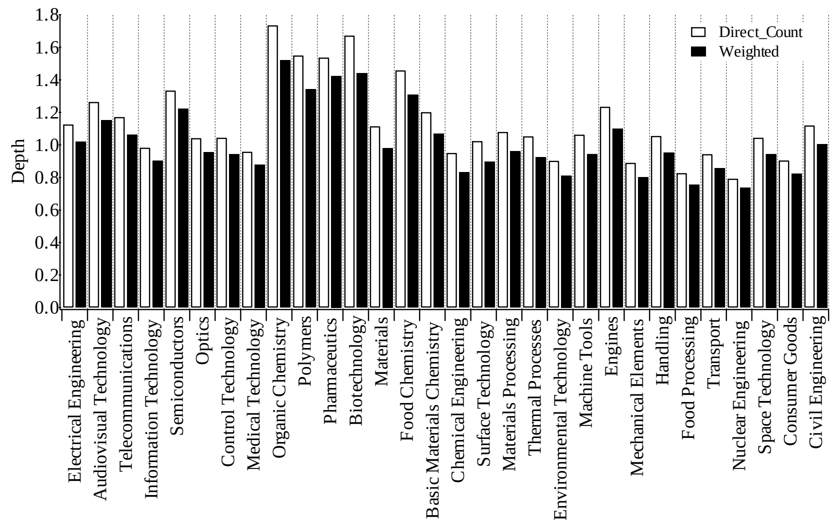

Figure 3: Weighted and Direct Count Depth Measures for 30 Main Technology Fields

IPC list of the patent. These depth measures are used as weights in the number of technology fields in the same area as the patent, given above by d_{ij}. In other words, one is concerned not only with the extent to which the patent draws upon knowledge in its own technology field, but also the extent to which a certain subject matter is utilized in depth by the patent. The final depth measure for technology field j is given by;

$$D_j = \frac{\sum d_{ij}(1 - w_{ij})}{N_j}$$

The Figure 3 demonstrates weighted and non-weighted depth measures for each technology field.

As can be seen in the figure, there are significant differences among technology fields in terms of the depth of patents, and the relative depth measures are robust with respect to the two measurement techniques.

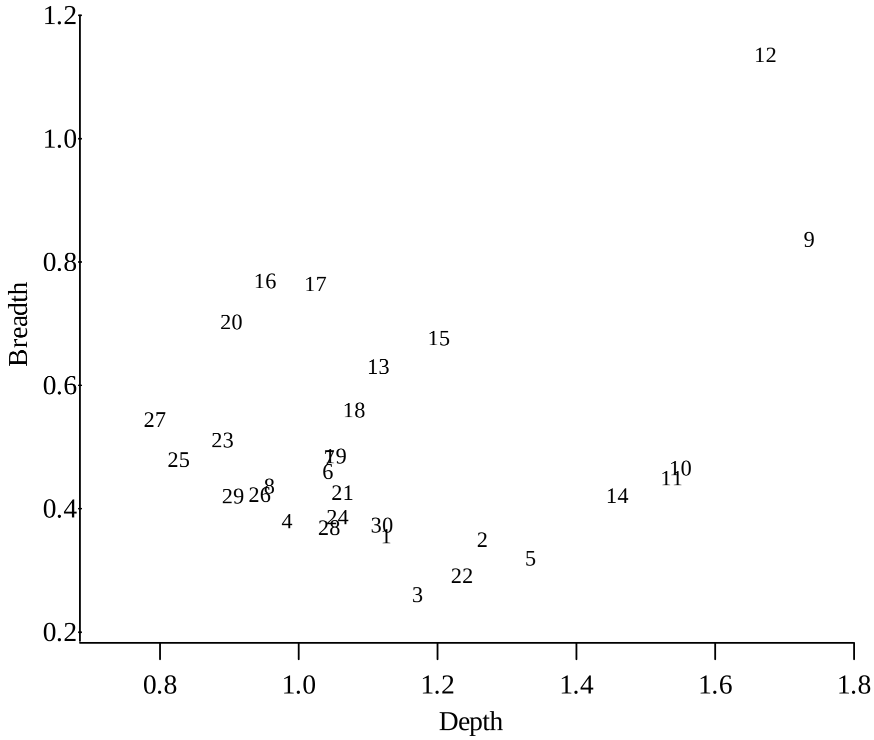

Figure 4: Breadth and Depth of 30 Main Technology Fields- Direct Count

2.5 Breadth and Depth

Figures 4 and 5 show the breadth and depth of the 30 technology fields by direct count and weighed techniques respectively (see Table 1 for descriptions of technology fields). It can be seen that the results in both methods are very robust, and the relative position of technology fields preserve their position in the breadth and depth space. Biotechnology is markedly different from other fields in terms of having the highest breadth and depth measures.

It is possible to see in these figures that technology fields are significantly different from each other with respect to their breadth and depth measures. Few observations stand-out from these figures. The results seem to be very robust to the measure used which confirms the relative position of technology fields with respect to each other in the breadth and depth space whether direct counting or weighting is used.

Table 1: List of Technology Fields

1	Electrical Engineering
2	Audiovisual Technology
3	Telecommunications
4	Information Technology
5	Semiconductors
6	Optics
7	Control Technology
8	Medical Technology
9	Organic Chemistry
10	Polymers
11	Pharmaceuticals
12	Biotechnology
13	Materials
14	Food Chemistry
15	Basic Materials Chemistry
16	Chemical Technology
17	Surface Technology
18	Materials Processing
19	Thermal Processes
20	Environmental Technology
21	Machine Tools
22	Engines
23	Mechanical Elements
24	Handling
25	Food Processing
26	Transport
27	Nuclear Engineering
28	Space Technology
29	Consumer Goods
30	Civil Engineering

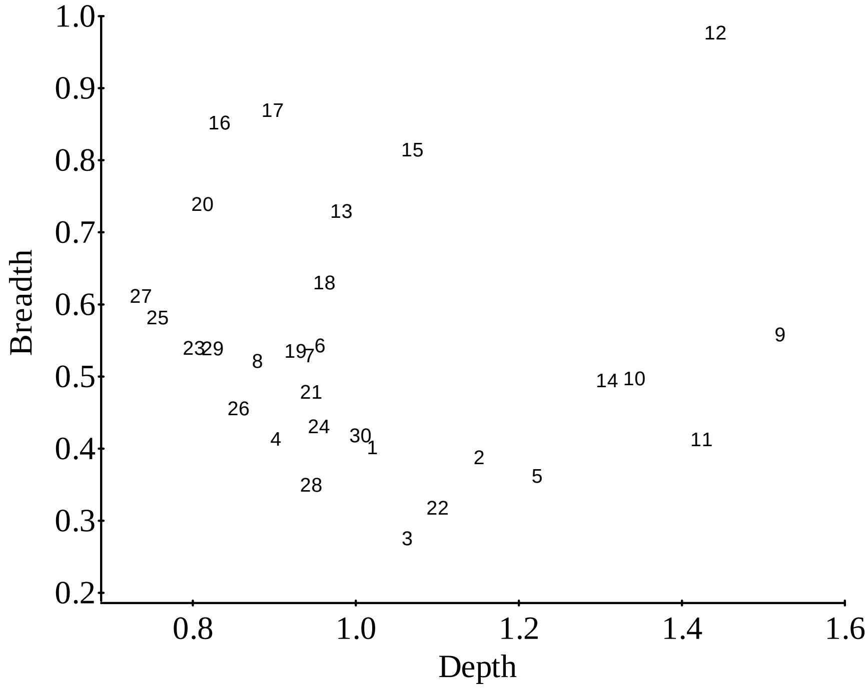

Figure 5: Breadth and Depth of 30 Main Technology Fields, Weighted

Biotechnology and organic chemistry stand out in terms of highest breadth and depth. Telecoms and semiconductors have very low breadth measures.

2.6 Evolution of Breadth and Depth 1978-2000 (30 Main Technology Fields)

Figures 6 and 7 present the evolution of breadth and depth in time for selected technology fields.

A general pattern is that for most fields, the breadth measures remains more or less stable, with the exceptions of telecoms, biotechnology and organic chemistry. For the telecoms, the knowledge base seems to have become more focused, while the opposite holds true for biotechnology and organic chemistry where the knowledge base expanded. For the depth measures, it is possible to see that nearly in all fields, the depth increases, which maybe explained by the increased specialization in most knowledge-intensive sectors.

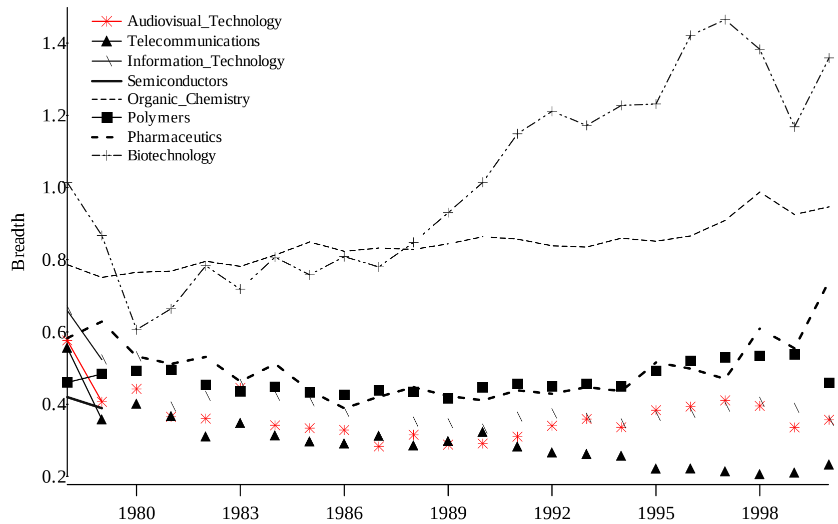

Figure 6: Evolution of Breadth in Selected High Tech Fields

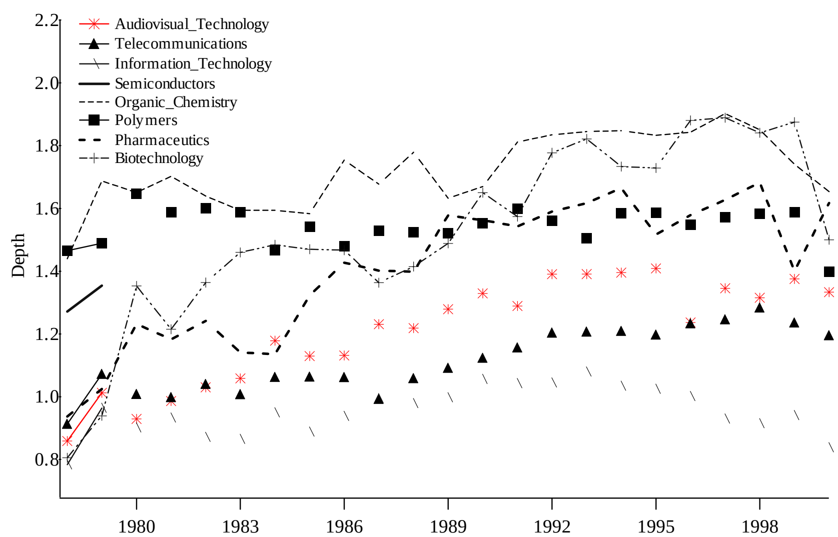

Figure 7: Evolution of Depth in Selected High Tech Fields

2.7 Breadth and Depth of Largest Firms

In this section, an analysis of the patents held by the largest firms in biotechnology and telecoms is performed. For this purpose, the patents are assigned to the firms, and the average of breadth and depth measure for the largest firms in biotechnology and telecoms are calculated. Formally, if patent i is taken by firm j, and if the corresponding breadth and depth of patent i are b_{ij} and d_{ij}, the breadth and depth of firm j is given by:

$$b_j = \frac{\sum_i b_{ij}}{n_j} \tag{1}$$

$$d_j = \frac{\sum_i d_{ij}}{n_j} \tag{2}$$

where n_j is the total number of patents taken by firm j in the period 1978-2000. The Figures 8 and 9 show our results for the two sectors biotechnology and telecoms.

3 Conclusion

In this paper, a methodology to measure the breadth and depth dimensions of knowledge bases is proposed. Breadth refers to the range of knowledge types that are used in the technology field. Depth refers to the extent to which a few knowledge types are exploited in depth. It was demonstrated that, the 30 main technology fields are largely scattered in the breadth and depth space. An important observation is concerning the biotechnology sector, where both breadth and depth of the field seems to be very high compared to other fields. The widening and deepening of the biotechnology knowledge base seems to have accelerated particularly in the last 2 decades. It was shown in this paper that not only technology fields, but also firms are also scattered in the breadth and depth space in terms of the patents they have taken.

These results have implications for innovation studies, such that recently it has been shown by many scholars that characteristics of the knowledge bases strongly influence innovative performance and/or organizational structures. Empirical research

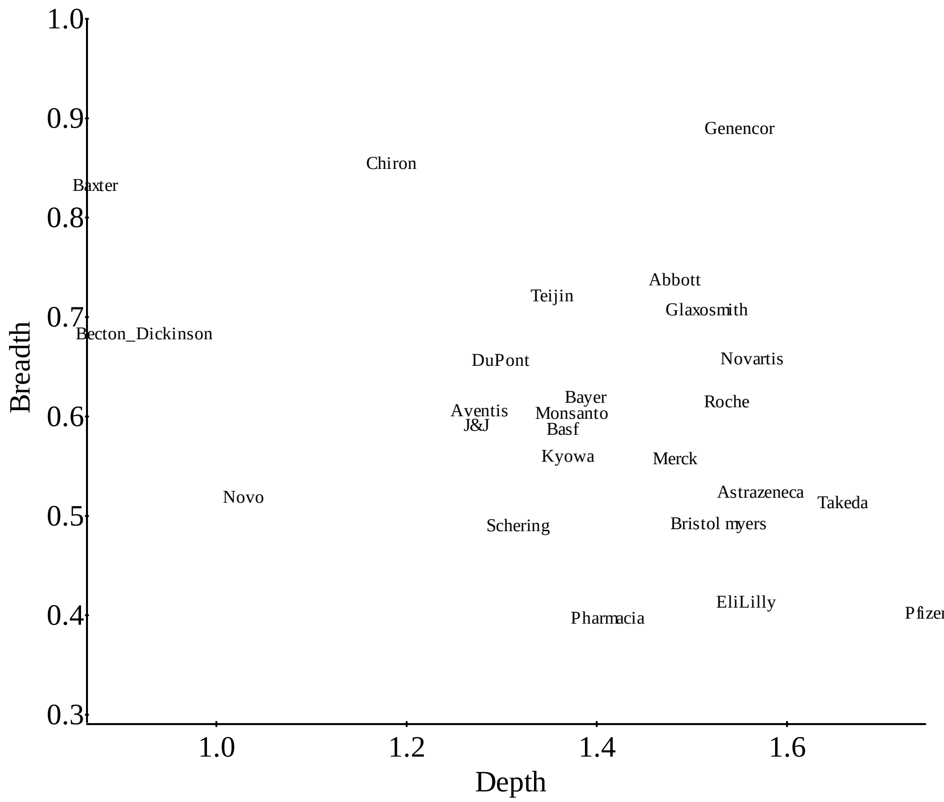

Figure 8: Breadth and Depth Measures of the Largest Biotechnology Firms

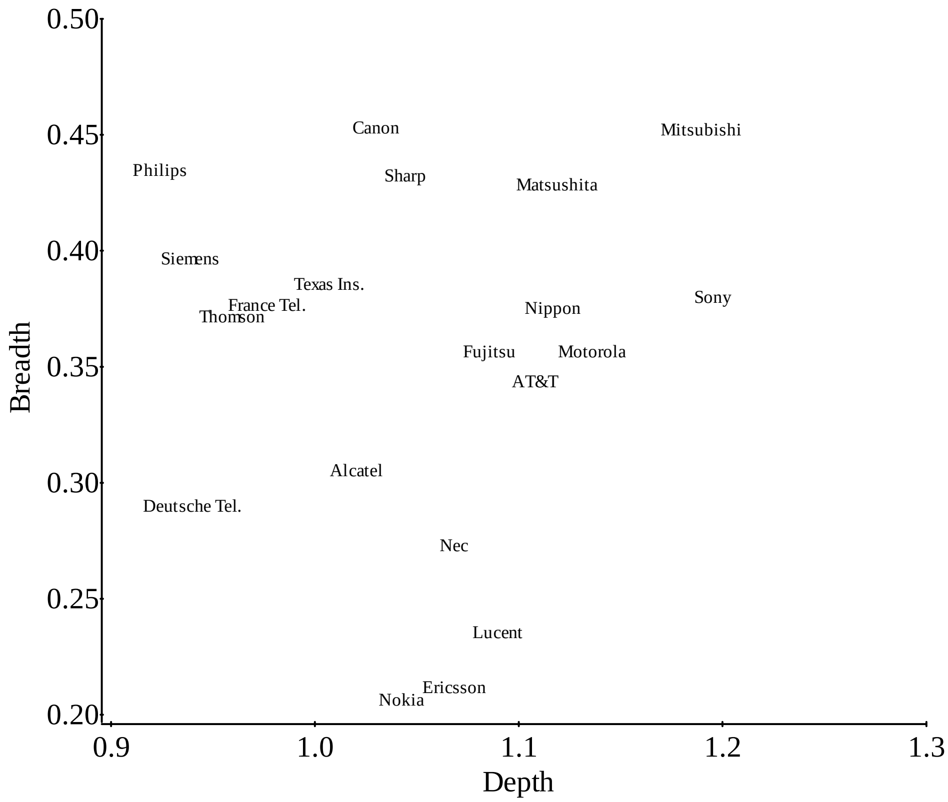

Figure 9: Breadth and Depth Measures of the Largest Telecommunications Firms

to measure characteristics of knowledge bases are is very limited, and in this paper, it is shown that technology fields do exhibit differences in terms of their knowledge bases. In this sense, this paper opens up new possibilities for future research to explore the contribution of breadth and depth to organizational structures and innovation performance.

Acknowledgement 1 *This research has been carried out using the European Patent Office database provided by Centre of Research on Innovation and Internationalization (CESPRI) in Bocconi University, Milano. I thank Franco Malerba and Stefano Breschi for many helpful discussions in constructing the breadth and depth measures, and for their help on the use of the database.*

References

[1] Breschi, S., Lissoni, F. and Malerba, F., 2003. Knowledge Relatedness in Firm Technological Diversification. Research Policy, 32: 69-87.

[2] Brusoni, S., Pavitt, K. and Prencipe, A., 2001. Knowledge Specialisation, Organisational Coupling, and the Boundaries of the Firm: Why Do Firms Know More Than They Make?. Administrative Science Quarterly, 26(4): 597-621.

[3] Brusoni, S. and Prencipe, A., 2000. Unpacking the Black Box of Modularity: Technologies, Products, Organizations. Industrial and Corporate Change, 10: 179-205.

[4] Cowan, R., N. Jonard and Ozman, M. 2003. Knowledge Dynamics in a Network Industry. Technological Forecasting and Social Change, 71(5): 469-84.

[5] Dosi, G and Hobday, M., 1999. Problem-solving Behaviour, Organisational Form and the Complexity of Tasks. Paper prepared for the Dynacom Project (European Union, DGXII, TSER) in collaboration with the UK Complex Product Systems Innovation Centre, funded by the ESRC. Pisa/Brighton: St Anna School of Advanced Studies/SPRU.

[6] Kauffman, S., 1993. The Origins of Order, New York: Oxford University.

[7] Nesta, L. and Saviotti, P., 2004. Coherence of the Knowledge Base and Firm Innovative Performance: Evidence from the US Pharmaceutical Industry", SPRU Electronic Working Paper Series, 113.

[8] Orsenigo, L., Pammoli, F. , Riccaboni, M., Bonaccorsi, A. and Turchetti, G., 1998. The Evolution of Knowledge and the Dynamics of an Industry Network. Journal of Management and Governance, 1:147-175.

[9] Ozman, M., 2005. Networks, Organizations and Knowledge. PhD thesis, MERIT, Maastricht University.

[10] Prencipe, A., 2000. Breadth and Depth of Technological Capabilities in the CoPS: The Case of the Aircraft Engine Control System. Research Policy, 29: 895-911.

[11] Simon, H. A., 1969. The Architecture of Complexity, in The Sciences of the Artificial. Cambridge, MA: MIT Press.

[12] Wang, Q., and Von Tunzelmann, 2000. Complexity and the Functions of the Firm: Breadth and Depth. Research Policy, 29: 805-818.

[13] Zander, U. and Kogut, B.,1995. Knowledge and the Speed of the Transfer and Imitation of Organizational Capabilities: An Empirical Test. Organization Science, 6(1): 76-92.