**Developing technology in the vicinity of science:
An examination of the relationship between science intensity (of patents) and technological
productivity within the field of biotechnology**

Bart Van Looy
Tom Magerman
Koenraad Debackere

Steunpunt O&O Statistieken
Research Division INCENTIM
Faculty of Economics and Applied Economics
K.U.Leuven

Naamsestraat 69
3000 Leuven
Belgium

Contact: Bart.vanlooy@econ.kuleuven.ac.be

Abstract

In this paper we investigate - on a country level - the relationship between the science intensity of patents and technological productivity, taking into account differences in terms of scientific productivity. The number of non patent references in patents is considered as an approximation of the science intensity of technology whereas a country's technological and scientific performance is measured in terms of productivity (i.e. number of patents and publications per capita). We use USPTO patent-data pertaining to biotechnology for 20 countries covering the time period 1992-1999. Our findings reveal mutual positive relationships between scientific and technological productivity for the respective countries involved. At the same time technological productivity is associated positively with the science intensity of patents. These results are confirmed when introducing time effects. These observations corroborate the construct validity of science intensity as a distinctive indicator and suggest its usefulness for assessing science and technology dynamics.

Acknowledgements:

The authors gratefully acknowledge financial support of the Flemish Community (Steunpunt O&O Statistieken) while conducting this research.

Introduction

There is a long history of using patent data to understand and assess invention and innovation dynamics in both economics and science and technology studies (e.g. Schmookler, 1966, Griliches, 1990, Jaffe, 1986). While patent based indicators indeed have limitations – not all inventions are being patented, while at the same time the propensity to patent differs among industries - addressing issues of (sources of) economic growth, the competitive position of countries or companies and the dynamism of industrial structures, requires a profound insight into patterns of (differential) inventiveness. In this respect patent databases and the indicators derived from it, are still one of the prime information sources given their coverage, transparency and accessibility. Griliches' observations - when surveying the state of the art on patent statistics 15 years ago - still seem to hold: "in this desert of data, patent statistics loom up as a mirage of wonderful plentitude and objectivity" (Griliches, 1990, p. 1661).

Not only have patent indicators become more and more institutionalized (see for instance OECD 2001, NSF, EC Science and Technology Indicator Report), over the last decade one has witnessed refinements in terms of the use of patent indicators to assess technological activities and to examine relationships with (economical) performance. Important evolutions relate to the use of (patent) citations to differentiate the value of patents (Trajtenberg, 2002, Harhoff, Scherer & Vopel, 2003) as well analyzing and mapping non-patent references –most of which are references to the scientific literature - found within patent documents (Narin & Noma, 1986, Collins & Wyatt, 1988, Van Vianen, Moed & Van Raan, 1990, Narin & Olivastro, 1992, Schmoch, 1993, Narin, Hamilton & Olivastro, 1997; McMillan, Narin & Deeds, 2000). These latter approaches can be related directly to the rising popularity of integrative notions like (1) scientific networks (Steinmueller, 1994; David, Foray and Steinmueller, 1997; Pavitt, 1997;), (2) strategy and its concomitant structural analysis of industries and competitors (Porter, 1995), (3) evolutionary economic thinking (Nelson, 1995) and (4) a new vision on industry, academia and government interactions as encompassed by the 'Triple Helix' model (Etzkowitz & Leydesdorff, 1997, 1998; Leydesdorff & Etzkowitz, 1996; 1998). Innovation effectiveness is conceived to an ever larger extent as stemming from (networks of) interactions unfolding amongst a variety of actors. These include companies as well as knowledge generating institutes like universities, while government agencies enact framework conditions in which such interactions can evolve effectively.

Stated otherwise, the interplay between technological and scientific realms is increasingly considered as essential for being effective in terms of knowledge creation, technology development and its translation into economical activity. This phenomenon of 'scientification' of technology development seems to hold particularly true for new emerging, knowledge intensive fields of economic activity (Gibbons et al.; 1994; Meyer-Krahmer, 2000; Tijssen, 2001).

These conceptions of innovation dynamics almost naturally led to efforts to delineate and develop indicators that – at least partially – grasp the complex set of interactions between both activity realms. Among the potential candidates, the nature of references found in patent documents has received considerable attention. Whereas the pioneering work of Jaffe, Trajtenberg and Henderson (1993; Jaffe A. and Trajtenberg M (1996) for an overview, see Jaffe and Trajtenberg, 2002) focused on the role of *patent* references and citations to map and assess knowledge spillovers and as an indication of the value of patents, Narin and his colleagues pioneered the role and possible contribution of non patent references. Studies in this field have been investigating the nature of the science-technology interaction that is implied by a citation link (e.g. Narin & Noma, 1985) the role of public science for developing technology (e.g. Narin et al. 1997), or still the frequency and nature of occurrence of such interactions in new emerging technology domains (Van Vianen et al. 1990; Meyer, 2000a, McMillan et al., 2000, Verbeek et al. 2002b).

Studies addressing the relationship between science intensity of patents – as measured by the amount of non patent references - on the one hand and the effectiveness of technology development on the other hand are less frequent to be found. On a country level, Van Looy et al. (2003) examined the relationship between the science intensity of patents and technological performance (productivity, revealed technological advantage). Positive relationships have been observed for so-called high tech domains like biotechnology, pharmaceuticals, organic fine chemistry and semiconductors while for other domains - including chemistry, food chemistry, measurement and control technology but also telecommunications – weak relationship were found. These observations led the authors to conclude that the relevancy of science intensity when developing technology is a domain specific phenomenon (see in this respect also Narin et.al. 1992; Grupp & Schmoch, 1992).

At the same time it can be observed that within the aforementioned analysis, indicators pertaining to the scientific capabilities of a country have not been taken into account. As such, the observed positive relationships observed might stem from the presence of scientific capabilities; in this case one would merely be counting ‘spill over’ effects that could be assessed equally by established bibliometric indicators pertaining to scientific publications. Hence, in order to assess the relevancy of using non patent references as an (additional) indicator to explain differences in technological performance, further analysis - whereby scientific capabilities are taken into account – is required. It is in this area that we want to situate our contribution. By Within the biotechnology domain – identified as a field in which the relationship between science and technology is intimate (see for instance, McMillan et al, 2000 ,Van Looy et al. 2003) - the following research questions are being addressed:

- Does the science intensity of patents – as measured by the amount of non patent references – relates to technological performance when scientific capabilities are brought into the equation?
- If the relationship between scientific capabilities and technological activity over time is to be conceived as bi-directional, what role does science intensity play in this respect?

Within this paper we analyze biotech patents granted within the USPTO patent system covering the time period 1992- 2002. In line with the concept of national innovation systems, countries are acting as the unit of analysis (Nelson, 1993). In the next sections we discuss the concepts and indicators used within this analysis and the findings obtained. However, before turning to the empirical analysis we first discuss in more detail the nature of non patent references found within patents as these references play a central role in the analysis undertaken.

A closer look at Non Patent References

As became clear in the introduction, recently considerable attention has been paid to the analysis of non patent references. At the same time, some concerns about the exact meaning of these references have been uttered. Meyer (2000, 2001), based on a limited number of patent cases studies, concludes that non patent references should not be interpreted as signaling a direct – and uni-directional - link or influence from science to technology as sometimes suggested by the rhetoric of advocates who depict non patent references as an indicator of science-technology interactions.

Tijssen (2001) points in a similar direction: non patent references should not be seen as reflecting scientific sources leading directly to the invention, but rather be considered a general indicator of ‘interaction’ between science and technology’ (Tijssen, 2001, p. 39). A closer look at the specific role references to prior art play within the patent application is in this respect highly informative. Within the next paragraphs we focus on the USPTO legislation and procedures as the data used within this analysis pertain to USPTO. For an exhaustive overview with an emphasis on EPO, we refer to Michel & Bettels (2001). At the same time, it can be noticed that the main difference between both systems with respect to citing prior art relates to the amount of references to be found rather than its nature; due to different disclosure obligations imposed on applicants.

Patents are documents issued by an authorized governmental agency which grants the right to the applicant to prevent others to produce or use a specific new device, apparatus or process for a limited time period. Patents are granted to the applicant after an examination that focuses on the novelty, inventive activity and industrial applicability. During the granting process, patent examiners review the prior art pertaining to the invention. While applicants are obliged – within the USPTO granting process - to provide an overview of all known relevant prior art, which can be either patents or other written documents, patent examiners do not limit themselves to the prior art signaled by inventors and/or applicants. Based on information, archives and databases available, patent examiners in the end decide which references are relevant to assess the claims made. Stated otherwise, it are the examiners references, used to decide on granting, including restricting claims, that are to be found on the front page of patent documents, besides information pertaining to the invention, assignee(s) and inventor(s). These references do not necessarily coincide with references provided by the applicant; references provided by the applicant can be omitted while examiners might add references as well.

Stated otherwise, front page references as found in patent documents are being introduced during the granting process for the purpose of evaluating the novelty and inventiveness of the claims and their applicability, including contextualizing the claims that are being made. As can be read in the patent examining procedure manual “The basic purpose for citing prior art in patent files is to inform

the patent owner and the public in general that such patents or printed publications are in existence *and should be considered when evaluating the validity of the patent claims*.

Placement of citations in the patent file along with copies of the cited prior art will also ensure consideration thereof during any subsequent reissue or re-examination proceeding.” (Manual of Patent Examining Procedure, USPTO, italics added; see also, 35 U.S.C. 301 and 37 CFR 1.501

It is clear that the specific role of references within patent application procedures is different from the role references or citations play within scientific publications. Within articles references indicate sources of influence or serve as reference points to delineate differences (novelty). At the same time, references to previous work in scientific publications are introduced by the authors (sometimes with some support of reviewers), implying in all cases that the cited references are known to the author(s) and hence have had a certain influence on the genesis of the ideas and insights developed within the article or paper. This clearly is not necessarily the case for the front page references to be found within patent documents. References might be added by examiners without the inventors and/or applicants being aware of their presence or without this knowledge having influenced in any way the creation of the invention, as documented recently by Meyer (2000) and Tijssen et al. (1999).

Against this background, a citation is perceived here as a bit of information linking two different documents. The presence of scientific research in the ‘prior art’ description of a patented invention, is seen as an indicator of the ‘distance’ between scientific findings on the one hand and technology development on the other hand. As references to be found in patents are a reflection of prior art, more references towards science fields signal more relevant prior art derived from scientific sources. While this does not equal a uni-directional, influencing or contributing, link from the cited paper towards the citing patent, it is clear that the more scientific references are considered relevant for assessing and contextualizing the claims made within the patent, the closer the technology is situated to scientific activity. Rather than equaling non patent references as signaling direct or uni-directional influences, contributing to the genesis and development of the invention at hand, they are part of the context in which the patent and its claims are to be situated. Hence, the more scientific references are to be found within patents, the more technology development is considered here as situated in the neighborhood or vicinity of scientific developments.

Indicators reflecting the amount of non patent references can be grasped through directly available and accessible data: sources. More specifically, we use the amount of non patent references as found within published USPTO patent documents as an indicator of the science intensity or science proximity of patents. A method towards the construction of this indicator was designed and implemented previously (see Verbeek et al. 2002a). The designed method encompasses a complex parsing algorithm, based on a textual analysis approach that parses and the references to be found in patents. Within the framework of the analysis reported here, the number of other references identified by means of this parsing algorithm is used as an indicator of the science intensity of each patent¹.

Data sources and indicators used.

Delineating the Biotechnology Domain

Biotechnology can be labeled as an industry in which the interplay between science and technology is important. From the seventies onwards scientific findings have been playing an important role within the industry (McMillan et al., 2000) resulting as well in numerous studies focusing on the role of collaboration and networking (Deeds et al. 1996; Baum et al. 2000; Rothaermel & Baum, 2004), including science-technology linkage which is particularly strong in this field (e.g. Narin and Noma, 1985, Murray, 2002; Verbeek et al., 2002;).

In order to select relevant patents we build further on previous research efforts focusing on the biotechnology industry. In order to retrieve all relevant biotechnology patents, we took the recent OECD definition of the biotechnology area which draws on five different patent classes of the

¹ A detailed content analysis of 10.000 other references reveals that about 60 % of these references are references towards scientific journals (See Callaert et al. , forthcoming). The remainder relates to books, company reports databases and the like. At the same time, the correlation between the number of other references and the number of journal references nearly equals 1; allowing to use the frequency of occurrence of other references to address the research questions outlined.

International Patent Classification as a starting point (OECD, STI Scoreboard 2001: 32; see also van Beuzekom, 2001)². In a next step we added two subclasses which the Fraunhofer Classification Scheme includes as biotechnology-relevant: C07G (*Compounds of unknown constitution*) and C12R (*indexing scheme related to subclasses C12C to C12Q or C12S, related to micro-organisms*). In addition to the International Patent Classification scheme, the US Patent and Trademark Office (USPTO) uses its own classification scheme. Based on the US classification, Jaffe and his colleagues at NBER set up an alternative classification scheme to the IPC-based Fraunhofer classification (e.g. Hall et al., 2001). Also their classification allows identifying biotechnology related patents. In accordance with this NBER classification scheme, we selected the US patent classes 435 and 800 to delineate biotechnology from other US patents. Equipped with these datasets, validation interviews were carried out with a number of field experts confirming the overall validity of the approach. As for publications, data retrieval was based on 9 relevant subject categories provided within the framework of the SCI Expanded database (ISI). Annex 1 provides a more elaborated description of the development of the search keys used. For the time period 1992 – 2001, this resulted in a total number of 51.460 granted patents and 967.188 publications.

Technological and Scientific Productivity

As an indicator of technological performance we used technological productivity³. In order to obtain this indicator, the yearly number of patents granted by country during the time period 1992-1999 was divided by the total population.

As data were retrieved in 2002, patents granted during the time period 2000-2001 have been omitted from the analysis as for these years the diminishing number of granted patents – due to the time delays implied by the granting process – might affect the relationships examined considerably. In this respect it can be remarked that the concept of technological productivity would be more accurately depicted by dividing the number of patents by the total amount of R&D expenditures within the field. While this is undoubtedly the case, the variety of disciplines and industries involved as well as the lack of reliable, distributed data – on a country level – on R&D expenditures that are to be attributed unambiguously to biotechnology activities prevented such an approach. For the analysis reported in the following section, logarithmic transformation of the productivity variable has been applied in order to obtain a normal distribution.

All patents are grouped by country based on the nationality of the patent assignee. In the case of co-patenting involving multiple countries, full counts were applied for all countries involved. Whereas patent data retrieved relate to patents granted during the time period 1992 – 1999; within the analysis we will use the application year of the patent data. Given that the time period between applying and granting averages between 2 and 3 years, application dates are preferred to assess the relationship with scientific capabilities. With respect to this latter, the total number of scientific publications – as retrieved from ISI Web of science – within biotechnology has been aggregated by country. Also here, in the case of co-authored papers implying different countries full counts have been applied. In order to examine the relationship between technological performance and scientific capabilities, patent indicators reflecting technology activity at moment T0 (application date) are related to scientific capabilities as measured at T1 to account for the delays observed in publications⁴.

Science Intensity

² C12M (Apparatus for enzymology or microbiology); C12N (Micro-organisms or enzymes; propagating, preserving, or maintaining micro-organisms; mutation or genetic engineering; culture media); C12P (Fermentation or enzyme-using processes to synthesise a desired chemical compound or composition or to separate optical isomers from a racemic mixture); C12Q (Measuring or testing processes involving enzymes or micro-organisms; compositions or test papers therefore; processes of preparing such compositions; condition-responsive control in microbiological or enzymological processes); and C12S (Processes using enzymes or micro-organisms to liberate, separate or purify a pre-existing compound or composition; processes using enzymes or micro-organisms to treat textiles or to clean solid surfaces of materials).

³ At the same time, we conducted a parallel analysis using technological performance as measured by the total number of patents (logarithmically transformed) and using population figures as independent variables. Such an analysis reveals similar results as obtained here with respect to the relationships between technological productivity, scientific productivity and science intensity (see Annex 2).

⁴ A analysis with a time lag of two years yielded similar results as the ones reported here.

In order to assess the science intensity of patents, the number of citations retrieved in patent documents is used here as an indicator. As became clear above and similar to the majority of studies that investigate the link between science and technology, we make use of so-called front-page references (United States Patent Trademark Office - USPTO). While it could be argued that such an analysis might benefit from including applicant given references (see for instance Meyer, 2000b), their present unavailability within the database does not allow for inclusion⁵. Hence, the analysis includes all examiner given references – to be found on the front page – which include in most cases a considerable amount of applicant references. For each country and each year, the average amount of references found within the assigned patents has been calculated and included within the analysis.

Results

Applying the search keys for patents and publications resulted in a dataset covering 20 countries⁶ demonstrating technology and scientific activities on a yearly base during the time period covered (1992 – 1999). Table 1 provides an overview of summary statistics related to the key indicators by country. Science Intensity reflects the average number of other references found within patents; scientific and technological productivity is expressed in terms of number of publications and patents per million inhabitants. For technology productivity a logarithmic transformation was performed in order to become a normal distribution.

Insert Table 1

Table 2 complements these data by providing an overview of the correlations observed between the variables. One observes positive, significant relationships between technological and scientific productivity; at the same time science intensity correlates positively with both productivity indicators.

Insert Table 2

Within a next step, partial correlations have been calculated relating the different key constructs one by one controlling for application year and the third variable. Table 3 summarizes the results.

⁵ Whether or not big differences would result from using either source remains to a large extent unclear. While the social processes in which applicant and examiner's roles are embedded might justify expected differences, empirical work that demonstrates these differences remains scarce. A detailed analysis of citations made in a sample of 366 patents in the genetic field (time period 1980-1985) by Collins and Wyatt (1988) revealed no major differences with respect to citations given by examiners and applicants. Also the recent analysis of Meyer (2000) indicates that the majority of applicant references tend to be included in the references assigned by examiners.

⁶ In total Biotechnology patents are found for over 50 countries; within this analysis we only used country data if patent activity was to be observed throughout the whole time period considered (1992-1999).

While one observes again positive relationship between scientific and technological productivity as well as science intensity and technological productivity, the positive relationship between science intensity and scientific productivity found without controlling for technological productivity disappears. Stated otherwise, the correlation observed between scientific productivity and science intensity can be attributed to the relationship between both variables and technological productivity.

Insert Table 3

The positive relationship between technological productivity on the one hand and scientific productivity and science intensity can be observed as well when performing a multiple regression analysis as table 4 makes clear. Both scientific productivity and science intensity turn out to be positively related to technological productivity.

Insert Table 4

These findings allow answering the first two research questions in an affirmative way: science intensity is positively associated with technological performance. This positive relationship is to be found as well when bringing scientific capabilities into the equation. Stated otherwise, countries in which patents include more other references display higher levels of technological activity. Notice that the same relationship holds for scientific capabilities: higher (lower) levels of scientific productivity coincide with higher (lower) levels of technological productivity.

Within a final analysis, we explore the relationship over time between the three variables under study. In a first step, additional variables were introduced reflecting technological and scientific productivity as well as science intensity at $T+2$ ⁷. Given the limited number of variables, we calculated the partial correlation coefficients between the variables in line with the path analysis logic outlined by Blalock (1961; see also Davis, 1985). Figure 1 summarizes the obtained partial correlation coefficients whereby only significant relationships are depicted⁸. Two observations are emerging: first of all evidence is found for the mutual or bi-directional influence of scientific and technological productivity over time.

Whereas technological productivity at $T+2$ is largely associated with past technological productivity (T), a positive and significant relationship with scientific productivity is observed and vice versa.

Second, the science intensity of patents not only seems to be a 'path dependent' phenomenon – as the positive relationship between science intensity at T and $T+2$ indicates – a distinctive and positive association with technological productivity ($T+2$) can be observed as well. Again, these findings reveal that the amount of non patent references coincides with varying levels of technological productivity. Stated otherwise; the more technology development is situated in the vicinity of scientific developments, the higher technological productivity. This relationship holds within the field of biotechnology - on a country level - even after introducing scientific productivity into the analysis.

Insert Figure 1

Discussion & directions for further research

Our findings reveal that within the field of biotechnology, the science intensity or science proximity – as measured by the amount of non patent references - of patents is positively associated with technological productivity. The relationship between science and technology within the field of biotechnology indeed reveals itself here as reciprocal and bi-directional rather than unidirectional or linear while at the same time both activity domains deploy their own 'internal' dynamics (Rip, 1992).

⁷ An analogous analysis has been conducted with $t+3$ variables, yielding similar results.

⁸ For the variables at $T+2$ all other variables (T and $T+2$) have been controlled for; while for variables at moment T , partial correlations imply a correction for the third variable (at moment T) only.

At the same time a distinctive relationship between science intensity and technological productivity has been observed. These findings corroborate the construct validity of indicators based on other references found within patents. In addition, the positive relationship between science intensity, or stated otherwise, the closeness between science and technology, and technological productivity, corroborates the relevancy of policy frameworks that foster interaction between knowledge/science generating institutions (universities, research centres) and technology producers (companies), at least in science intensive fields like biotechnology. At the same time these findings suggest interesting avenues for further research. While we focused on one specific field (biotechnology), refining the insights obtained in terms of their field specific nature requires extensions towards other fields. Likewise, introducing extended time frames would allow assessing whether differential effects are to be observed related to technological life cycle dynamics (Abernathy and Utterback, 1991; Andersen 2001).

References

- Andersen, B. (2001) Technological Change and the evolution of corporate innovation. Edward Elgar Publishing.
- Baum, J. A. C., Calabrese, T. and Silverman, B. S. (2000) Don't go it alone: alliance network composition and startups' performance in Canadian biotechnology. *Strategic Management Journal*, 21 (3): 267-294
- Blalock, H. (1961). Causal inferences in nonexperimental research. Chapel Hill, NC: UNC Press.
- Collins P. and Wyatt S. (1988) Citations in patents to the basic research literature. *Research Policy*, 17, 65-74.
- Callaert, J., Van Looy, B., Verbeek, A., Debackere, K. and Thijs, B. Traces of Prior Art: A systematic analysis of other references found within the USPTO and EPO patent system. Working Paper DTEW, K.U.Leuven.
- Davis, James A.. (1985). The logic of causal order. Quantitative applications in the social sciences series, no. 55. Thousand Oaks, CA: Sage Publications. 38 - 44 provide a non-technical introduction to partial correlation inferences
- Deeds, D. L. and Hill, C. W. (1996) Strategic alliances and the rate of new product development: an empirical study of entrepreneurial biotechnology firms. *Journal of Business Venturing*, 11: 41-55.
- De Solla Price, D. (1965) Is technology historically independent of Science? A study in statistical historiography. *Journal of Technology and Culture*, 6, 553-568.
- De Solla Price, D. (1963) Little Science, Big Science. Columbia University Press.
- Etzkowitz, H. and Leydesdorff, L. (1997). Introduction to special issue on science policy dimensions of the Triple Helix of university-industry-government relations. *Science and Public Policy*, Vol. 24, 1, 2-5.
- Etzkowitz, H. and Leydesdorff, L. (1998). The Endless Transition: A 'Triple Helix' of University-Industry-Government Relations. *Minerva*, 36, 203-208.
- Etzkowitz, H. & Leydesdorff, L. (2000) The dynamics of innovation: from National Systems and 'Mode 2' to a Triple Helix of university-industry-government relations. *Research Policy*, 29, 109-123.
- Griliches, Z. (1990) Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, 28, 1661 – 1707.
- Grupp, H. and Schmoch, U. (1992) Perception of Scientification of Innovation as Measured by Referencing between Patents and Papers. H. GRUPP (ed.) *Dynamics of Science-Based Innovations*, Springer Publishers, Berlin/Heidelberg, 73-128.
- Harhoff D, Scherer FM and Vopel K (2003) Citations, family size, opposition and the value of patent rights. *Research Policy*, 32, 8, 1343-1363.
- Jaffe A., Trajtenberg M, Henderson R. (1993) Geographical localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 108, 3, 577-598
- Jaffe A. and Trajtenberg M (1996) Flows of knowledge from universities and federal laboratories: Modelling the flow of patent citations over time and across institutional and geographic boundaries. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 23, 12671-12677.
- Jaffe A. and Trajtenberg M (2002) *Patents, Citations, and Innovations: A Window on the Knowledge Economy*. MIT Press.
- Leydesdorff, L. and Etzkowitz, H. (1996). Emergence of a Triple Helix of University-Industry-Government Relations. *Science and Public Policy* 23, 5, October 1996, 279-286.
- Leydesdorff, L. and Etzkowitz, H. (1998). Triple Helix of innovation: introduction. *Science and Public Policy*, Vol. 25, 6, 358-364.
- McCulloch, R. (1980) International indicators of science and technology – how does the US compare? *Scientometrics*, 2, 5-6, 355-367.
- McMillan, S., Narin, F. and Deeds, D. (2000) An analysis of the critical role of public science in innovation: the case of biotechnology. *Research Policy*, 29, 1-8.
- Meyer, M. (2000) Patent citations in a novel field of technology – What can they tell about interactions between emerging communities of science and technology. *Scientometrics*, 48, 2, 151-178.
- Meyer, M. (2000b) Does science push technology? Patents citing scientific literature. *Research Policy*, 29, 409-434.
- Meyer-Krahmer, F, Schmoch, U. (1998). Science-based technologies: university-industry interactions in four fields. *Research Policy*, 27, 835-851.
- Michel, J. and Bettels, B. (2001) Patent citation analysis: A closer look at the basic input data from patent search reports. *Scientometrics*, 51, 1, 185-201.

- Narin, F. and Noma, E. (1985) Is technology becoming science? *Scientometrics*, 7, 369-381.
- Narin, F. and Olivastro, D. (1992) Status report: Linkage between technology and science. *Research Policy*, 21, 237-249.
- Narin F. & Olivastro D. (1998) Linkage between patents and papers: an interim EPO/US comparison, *Scientometrics*, 41, pp.51-59.
- Narin, F., Hamilton, K. and Olivastro, D. (1997) The increasing linkage between US technology and public science. *Research Policy*, 317-330.
- Nelson, R.R. (1993) (Ed.) *National Innovation Systems: A Comparative Analysis*. Oxford University Press.
- OECD (1994), *The Measurement of scientific and technological activities: Using Patent Data as Science and Technology Indicators*. Patent Manual 1994, Paris.
- OECD (2001) *STI Scoreboard 2001*, OECD, Paris.
- Porter, M. (1995) *The Competitive Advantage of Nations*, New York. The Free Press.
- Rip, A. (1992) Science and technology as dancing partners. In Kroes, A. & Bakker, M. (eds.) *Technological Development and Science in the Industrial Age*. Kluwer Publications.
- Rothaermel, F. T. and Deeds, D. L. (2004) Exploration and exploitation alliances in biotechnology: a system of new product development. *Strategic Management Journal*, 25: 201-221.
- Schmoch, U. (1993) Tracing the knowledge transfer from science to technology as reflected in patent indicators. *Scientometrics*, 26, 1, 193-211.
- Schmoch, U. (1997) Indicators and the relations between Science and Technology. *Scientometrics*, No 38, 103-116.
- Soete, L.G and Wyatt, S.M.E. (1983). The Use of Foreign Patenting as an Internationally Comparable Science and Technology Output Indicator. *Scientometrics*, Vol.5, No.1, 31-54.
- Tijssen, R. (2001) Global and domestic utilization of industrial relevant science: patent citation analysis of science-technology interactions and knowledge flows. *Research Policy*, 30, 35-54.
- Tijssen, R.J.W., Buter, R.K and Van Leeuwen, Th.N. (2000). Technological relevance of science: validation and analysis of citation linkages between patents and research papers. *Scientometrics*, 47, 389-412.
- Tijssen, R.J.W. (2002) Science dependence of technologies: evidence from inventions and their inventors. *Research Policy*, 31, 4, 509-526.
- Van Beuzekom, B. (2001) *Biotechnology Statistics in the OECD Member Countries: Compendium of Existing National Statistics*. *STI Working Paper* 2001/6, 2001
- Van Looy, B., Zimmermann, E., Veugelers, R., Mello, J. and Debackere, K. (2003) Do Science-Technology interactions pay off? An exploratory investigation of 10 science intensive fields. *Scientometrics*, 57, 355-367
- Van Viaenen, B., Moed, H. and Van Raan, A.. (1990) An exploration of the science base of recent technology. *Research Policy*, 19, 61-81.
- Verbeek, A., Debackere, K., Luwel, M., Andries, P., Zimmermann, E., and Deleus, F. (2002a). Linking science to technology: Using bibliographic references in patents to build linkage schemes. *Scientometrics*, Vol. 54, No. 1-2.
- Verbeek, A., Callaert, J., Andries, P., Debackere, K., Luwel, M., and Veugelers, R. (2002b), *Science and Technology Interplay – A Modelling approach on a regional level*, Final Report to the EC DG Research, Brussels (also published in the EC Indicators report 2003)

Annex 1: Search Keys Patent. - Publications

Search Key Patents:

The starting point for our analysis was the recent OECD definition of the biotechnology area. It draws on the following patent classes of the International Patent Classification (OECD, STI Scoreboard 2001: 32; see also van Beuzekom, 2001):

- C12M (Apparatus for enzymology or microbiology);
- C12N (Micro-organisms or enzymes; propagating, preserving, or maintaining micro-organisms; mutation or genetic engineering; culture media);
- C12P (Fermentation or enzyme-using processes to synthesise a desired chemical compound or composition or to separate optical isomers from a racemic mixture);
- C12Q (Measuring or testing processes involving enzymes or micro-organisms; compositions or test papers therefore; processes of preparing such compositions; condition-responsive control in microbiological or enzymological processes);
- C12S (Processes using enzymes or micro-organisms to liberate, separate or purify a pre-existing compound or composition; processes using enzymes or micro-organisms to treat textiles or to clean solid surfaces of materials).

Furthermore, we added two subclasses which the Fraunhofer Classification Scheme includes as biotechnology-relevant: C07G (*Compounds of unknown constitution*) and C12R (*indexing scheme related to subclasses C12C to C12Q or C12S, related to micro-organisms*). Using the aforementioned IPC subclasses, a search for patent documents was carried out both in the US patent databases.

In addition to the International Patent Classification scheme, the US Patent and Trademark Office (USPTO) uses its own classification scheme. Based on the US classification, Jaffe and his colleagues at NBER set up an alternative classification scheme to the IPC-based Fraunhofer classification (e.g. Hall et al., 2001). Also their classification allows to identify biotechnology related patents. In accordance with this NBER classification scheme, we selected the US patent classes 435 and 800 to delineate biotechnology from other US patents. There are few differences between US patents retrieved through IPC classes and the set of US patents that was defined based on the US classification. As initial tests indicated, the IPC classification identified 8 % more patents than US patent classification.

Based on these datasets, validation interviews were carried out with a number of field experts from the Belgian and Flemish biotechnology industry and research. The interviews confirmed the validity of our overall approach in particular for fields of 'modern biotechnology'. All important actors in the biotechnology area were identified. However, the interviews also indicated certain fields were not covered to the extent the experts we consulted would have expected. In collaboration with them, we added additional IPC subclasses to our search strategy, in particular in the area of health and food-biotechnology. These areas are covered mainly by the sub classes A61K (health) and A23C (food).

In order to avoid the inclusion of non-biotechnology patents we used the IPC not only at the subclass but also at the group-levels. Table 1 gives an overview of the different search strategies in the biotechnology domain. We applied the broadest possible search strategy, encompassing patents after the OECD, Fraunhofer, and NBER classification as well as our additional classes and groups as identified in our validation exercise.

Table: Overview of search strategies for biotechnology patents used in different patent databases and studies

Strategy	Patent IPC
OECD definition on basis of IPC classification	C12M; C12N; C12P; C12Q and C12S
Fraunhofer classification	C07G; C12M; C12N; C12P; C12Q; C12R and C12S
NBER based US patent classification	435 and 800 (OCL classes)
Extended search strategy including modern and traditional biotechnology	A23C*; A23J*; A61K*; C07H; C07J; C07K; C07G; C12M; C12N; C12P; C12Q; C12R and C12S

From these classes a selection of subclasses were made as seen in Appendix A)

Search Key Publications

The delineation of the area was based on a set of journals assigned by the ISI to relevant subject categories. All papers published in these journals recorded as article, letter, note or review covered by the SCI Expanded have been retrieved, and have been added to our biotechnology database. For the macro analysis and for the determination of the most active institutions in the world, a set of nine Subject Categories has been selected which are assumed to form the core in biotechnology.

In particular, the following nine ISI Subject Categories have been used:

1. DE PLANT SCIENCES
2. CO BIOCHEMICAL RESEARCH METHODS
3. CQ BIOCHEMISTRY & MOLECULAR BIOLOGY
4. DA BIOPHYSICS
5. DB BIOTECHNOLOGY & APPLIED MICROBIOLOGY
6. QU MICROBIOLOGY
7. DR CELL BIOLOGY
8. KM GENETICS & HEREDITY
9. HY DEVELOPMENTAL BIOLOGY

Annex 2: Regression Model

Technological Performance (Log Patents) as dependent Variable
Scientific Capabilities, Science Intensity, Population, Application Year as Independent Variable

R	R Square	Adjusted R Square	Std. Error of the Estimate
,865	,749	,742	,34363

Coefficients					
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-11,180	24,839		-,450	,653
Application Year	4,218E-03	,012	,014	,340	,734
Science Intensity	1,277E-02	,004	,134	3,096	,002
Population	6,643E-03	,101	,005	,066	,948
Scientific Capabilities	1,206	,112	,833	10,798	,000

Table 1: Descriptive Statistics - Science Intensity, Technological Productivity & Scientific Productivity by Country

	Science Intensity			Scientific Productivity			Technological Productivity		
	Mean	Std Deviation	Valid N	Mean	Std Deviation	Valid N	Mean	Std Deviation	Valid N
Austria	14,81	4,16	N=8	101,82	21,17	N=8	1,13	,42	N=8
Australia	14,72	5,42	N=8	138,18	18,39	N=8	1,94	1,11	N=8
Belgium	11,32	3,56	N=8	145,30	19,13	N=8	2,10	1,78	N=8
Canada	24,12	4,90	N=8	150,59	5,09	N=8	3,08	1,23	N=8
Switzerland	11,10	3,39	N=8	277,56	28,51	N=8	5,24	1,74	N=8
Germany	11,32	2,38	N=8	101,70	13,90	N=8	1,95	,71	N=8
Denmark	10,25	3,12	N=8	232,76	26,06	N=8	10,80	4,90	N=8
Spain	8,17	5,30	N=8	64,55	27,93	N=8	,14	,08	N=8
Finland	22,47	8,31	N=8	149,40	63,08	N=8	2,34	,98	N=8
France	11,44	2,99	N=8	116,10	10,45	N=8	2,02	1,11	N=8
Great Britain	15,14	5,31	N=8	159,03	12,65	N=8	2,25	,88	N=8
Israel	18,81	8,19	N=8	197,35	13,13	N=8	3,90	1,19	N=8
Italy	5,98	2,70	N=8	67,30	8,01	N=8	,38	,19	N=8
Japan	6,65	1,06	N=8	78,30	8,01	N=8	2,82	,63	N=8
Korea	6,23	2,46	N=8	23,28	12,51	N=8	,39	,15	N=8
Netherlands	13,60	4,10	N=8	179,79	11,88	N=8	4,08	1,46	N=8
Norway	12,23	4,83	N=8	135,53	17,91	N=8	,99	,41	N=8
Sweden	17,22	8,66	N=8	255,94	24,30	N=8	2,31	1,03	N=8
Taiwan	12,65	7,99	N=8	31,43	6,43	N=8	,45	,29	N=8
United States	23,53	4,07	N=8	123,59	7,40	N=8	9,08	3,47	N=8
Total	13,59	7,15	N=160	136,48	70,39	N=160	2,87	3,12	N=160

Table 2: Correlations between Scientific Productivity, Technological Productivity and Science Intensity

		Application Year	Scientific Productivity	Technological Productivity	Science Intensity
Application Year	Pearson Correlation	1	,147	,110	-,041
	Sig. (2-tailed)	,	,065	,168	,609
	N	160	160	160	160
Scientific Productivity	Pearson Correlation	,147	1	,683**	,255**
	Sig. (2-tailed)	,065	,	,000	,001
	N	160	160	160	160
Technological Productivity	Pearson Correlation	,110	,683**	1	,399**
	Sig. (2-tailed)	,168	,000	,	,000
	N	160	160	160	160
Science Intensity	Pearson Correlation	-,041	,255**	,399**	1
	Sig. (2-tailed)	,609	,001	,000	,
	N	160	160	160	160

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

Table 3: Partial Correlations between Scientific Productivity, Technological Productivity and Science Intensity

		Scientific Productivity	Technological Productivity	Science Intensity
Scientific Productivity	Pearson Correlation	1	,647**	-0,0173
	Sig. (2-tailed)	,	,000	,829
	N	160	160	160
Technological Productivity	Pearson Correlation	,647**	1	,320**
	Sig. (2-tailed)	,000	,	,000
	N	160	160	160
Science Intensity	Pearson Correlation	-0,0173	,320**	1
	Sig. (2-tailed)	,0829	,000	,
	N	160	160	160

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

Table 4: Regression Model:
Dependent Variable: Technological Productivity
Scientific productivity, Science Intensity and Application Year as Independent Variable

Model Summary					
R		R Square	Adjusted R Square		Std. Error of the Estimate
,722		,521	,512		,35633
Coefficients					
	Un-standardized Coefficients	Std. Error	Standardized Coefficients	t	Sig.
(Constant)	B		Beta		
	-13,579	24,880		-,546	,586
Science Intensity	1,737E-02	,004	,243	4,232	,000
Application Year	6,490E-03	,012	,029	,520	,604
Scientific Productivity	4,468E-03	,000	,617	10,619	,000

Figure 1: Partial Correlation Coefficients
Technological and Scientific Productivity, Science Intensity – T0 and T+2