Data Mining and Text Mining for Science & Technology Research

Chapter · January 2005

DOI: 10.1007/1-4020-2755-9_9

Gerhard Paass
Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS

Handbook of Quantitative Science and Technology Research

Handbook of Quantitative Science and Technology Research

The Use of Publication and Patent Statistics in Studies of S&T Systems

edited by

Henk F. Moed

Centre for Science and Technology Studies,
University of Leiden, The Netherlands

Wolfgang Glänzel

Steunpunt O&O Statistieken,
K.U. Leuven, Belgium

and

Ulrich Schmoch

Fraunhofer Institute for Systems and Innovation Research,
Karlsruhe, Germany

KLUWER ACADEMIC PUBLISHERS
NEW YORK, BOSTON, DORDRECHT, LONDON, MOSCOW

eBook ISBN: 1-4020-2755-9
Print ISBN: 1-4020-2702-8

Created in the United States of America

Visit Springer's eBookstore at: http://ebooks.springerlink.com
and the Springer Global Website Online at: http://www.springeronline.com

Contents

Preface

This handbook offers a state-of-the-art overview of quantitative science and technology research. It focuses on the development and application of indicators derived from data on scientific or scholarly publications and patents. It comprises 34 chapters written by leading specialists in the various sub-domains. These chapters deal with theoretical and methodological issues, illustrate applications, and highlight their policy context and relevance. Authors present a survey of the research topics they address, and show their most recent achievements.

The 34 chapters are arranged into 5 parts: Disciplinary Approaches; General Methodology; The Science System; The Technology System; and The Science–Technology Interface. The Editor's Introduction provides a further specification of the handbook's scope and of the main topics addressed in its chapters.

This handbook aims at four distinct groups of readers:

– practitioners in the field of science and technology studies;
– research students in this field;
– scientists, scholars and technicians who are interested in a systematic, thorough analysis of their activities;
– policy makers and administrators who wish to be informed about the potentialities and limitations of the various approaches and about their results.

The current handbook can be considered as the successor of the Handbook of Quantitative Science and Technology Studies edited by Anthony van Raan and published in 1988 (Amsterdam: North-Holland).

We are most grateful to all contributors to the handbook for their enormous efforts to provide us with a series of excellent papers. We wish to thank Suze van der Luijt, Ed Noyons, and Renald Buter (CWTS) for their help in the technical editing process and in the preparation of the handbook's Subject Index.

Ulrich Schmoch gratefully acknowledges financial support for his editorial work from the German Federal Ministry for Education and Research (BMBF).

Henk F. Moed, Wolfgang Glänzel, and Ulrich Schmoch

EDITORS' INTRODUCTION

Henk F. Moed, Wolfgang Glänzel, and Ulrich Schmoch

This handbook deals with the quantitative study of the science and technology system from a global perspective. It provides a state of the art of the development and application of indicators for that system that are derived from publications, particularly — though not exclusively — from the scholarly literature and the patent literature. The science and technology (S&T) system comprises a wide range of activities from basic, or fundamental, science or scholarly activity, via strategic or application oriented research, to applied research and developmental activities aimed at the production of new products and processes. It can be conceived as a part of the various national or regional innovation systems.

During the course of the twentieth century, particularly after the Second World War, science and technology have become driving forces in society and vehicles of economic growth and development. The more important they became, the more the need was recognised to monitor their development, to examine the conditions under which they reach an optimal performance, and to formulate and carry out policies aimed at enhancing its performance and setting its priorities.

Science and technology policy and management — at the level of research group leaders, company technology managers, research programme managers, institutional directors, funding agencies, or at a regional, national or even supra-national political agencies — is itself scholarly based. In order to be effective, policy measures and decisions should be informed, and based upon proper insight into the functioning of the S&T system.

The contributions to this handbook reflect a wide variety of attributes of the S&T system that are relevant in such policies and numerous methodologies assessing such attributes. Central concepts are scientific or technological performance, and productivity or efficiency of the S&T system and its constituent parts. Crucial research questions are: how performance or

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 1-15.

productivity could be measured; how the various parts in an S&T system react one with another; how this interaction influences the overall performance; whether there are significant differences in performance amongst parts and how such differences can be explained.

The general issue that all studies deal with — most explicitly, some more implicitly — is the identification of factors or conditions which may positively or negatively effect the S&T system's performance defined broadly in terms of the needs and criteria expressed by the societies in which they are embedded. An overview of all contributions is presented in Table 1 at the end of this chapter.

This handbook presents analyses at the level of individuals, research groups, researcher networks, institutions, and at a regional, national and even supra-national level. Important attributes related in the various contributions to performance or productivity are: the availability of scientific–technological information; quality control mechanisms in knowledge production processes; internationalisation and globalisation; collaboration; knowledge networks and knowledge flows; multi- or interdisciplinarity; knowledge specialisation and integration; and participation of women.

In addition, several contributions discuss the wider policy and political context in which S&T indicators are actually used. Important aspects are criteria and conditions for a proper, informed use of indicators in performance assessments, and the possible effects that application of such indicators in assessments or funding procedures may have upon scholars and technicians subjected to such assessments.

As indicated above, this handbook primarily relates to indicators derived from scholarly publications and patents. A most important data source for analysis of the science system is the *Science Citation Index* (SCI) and related Citation Indexes published by the *Institute for Scientific Information* (ISI–Thomson Scientific, Philadelphia, PA, USA), or, in a more recent version, ISI's *Web of Science*. Once citation indexing became available for bibliographic research, it was apparent that they could be used to answer inquiries into the nature of scholarly activity: how it is structured; how it develops and how its actors perform. Garfield expressed this as follows:

> "If the literature of science reflects the activities of science, a comprehensive, multidisciplinary citation index can provide an interesting view of these activities. This view can shed some useful light on both the structure of science and the process of scientific development" (Garfield, 1979, p. 62).

As pointed out by Lionel Nesta and Pari Patel in this handbook, a patent is a legal instrument which confers a temporary monopoly of an invention in exchange for the publication of its details. Thus a patent has two functions: it protects the invention; but at the same time it disseminates knowledge about it. A patent can therefore be conceived as a publication which makes scientific–technological content public, similarly to a research article in a scientific journal. As discussed by Elise Bassecoulard and Michel Zitt in their contribution, scholarly publications and patents as formal information sources have many features in common. Similarly to the relation between publications and science, patents are used in the construction of indicators of trends, performance and structures in technological activity.

The development of S&T indicators takes place in various disciplinary contexts. The disciplinary approaches most prominently represented in this handbook use methods adopted from *physics*, *economic sciences*, *sociology*, *history of science and technology* and *information and communication science*, respectively. The first part of the handbook presents a number of contributions which illustrate such approaches, and may in this sense be seen as exemplifications.

A first disciplinary approach is that of *physics*. This approach has a long tradition in quantitative science and technology studies. Science and technology are conceived as a physical system of interacting sub-units the behaviour of which can be described by more general laws analogously to physical laws. In the first chapter ***Anthony van Raan*** reviews recent studies adopting this approach that are inspired by modern developments in the physics of non-linear phenomena. Van Raan also gives a thorough survey of main methodologies applied in the measurement of scientific activity.

A second disciplinary approach is that of *economics or econometrics*. It considers activities in the S&T system essentially as an economic activity in which the consumption of essentially scarce input resources leads to a number of identifiable scientific or technological ‘outputs’. Patenting is particularly conceived as an economic act. ***Andrea Bonaccorsi*** and ***Cinzia Daraio*** review and discuss the use of econometric methods in the study of the S&T system, and focus on the concept of productivity. They conceive S&T production as a non-deterministic, multi-input, multi-output relation, in which both inputs and outputs are not only qualitatively heterogeneous but also sometimes truly incommensurable.

Hariolf Grupp and ***Mary Mogee*** apply an economic policy approach and describe the increasing use of S&T indicators in the context of national policies, with a focus on the United States and Europe, and critically discuss the appropriateness of composite indicators, national benchmarking, and scoreboarding.

A third disciplinary approach is of *sociology* focussing either on social relationships and activities within the S&T system or on the relationships between this system and the wider socio-political environment. Activities of the various actors in the system are essentially conceived as social acts that are studied in their broader social context. This is also true for scholars' publication and referencing practices.

In order to be really informative, and particularly to constitute a sound basis for S&T indicators, scientific publications must meet professional standards. *Gatekeepers*, that is, the members of the editorial and advisory boards of science journals occupy powerful strategic positions in the collective activity of science. In a sociological study of the science system, ***Tibor Braun*** discusses characteristics of the journal gatekeeping system aimed at ensuring such standards.

Rémi Barré analyses the changing relationships between the S&T system and society in general, and highlights their implications for the role of science and technology indicators and their producers in S&T policy-making processes.

A fourth disciplinary approach is the *history of science and technology,* adopting a historical or evolutionary perspective. Bibliometric indicators can be used to trace developments in scholarly disciplines or technological areas and identify key events that can be used as reference points. As such, they are tools for a historian of science and technology for obtaining a historical account of such developments and for relating such developments to other factors from their wider socio-political or economic environment. The contribution by ***Birgitte Andersen*** represents an example of this approach. It aims to identify and measure changes in technological opportunities during the last century in order to trace the evaluation of their trajectories governed by technological 'paradigms'. This perspective is, of course, closely linked to evolutionary economics as well.

Information and communication science constitutes the fifth disciplinary approach represented in this handbook. It analyses how scholars or technicians in any field use and disseminate information through formal and informal channels, and identifies patterns or structures of the communication system. A bibliometric viewpoint thus focuses on scholarly or technical texts or documents. The contribution by ***Subbiah Arunachalam*** highlights the crucial importance of having access to up to date scientific and technical information, particularly for S&T practitioners in developing countries. In relation to this he presents an overview of current developments which could make access to information for scientists in those countries more affordable, including the emergence of open access journals and e-print archives.

The contributions in the subsequent parts of the handbook — although equally important and informative — apply elements from several disciplinary approaches rather than one, or use other disciplines not mentioned above, and structure these either within a particular methodology or within specific attributes or sub-systems of the S&T system. These contributions were grouped into four parts on the basis of their main emphasis: *methodological contributions*; and contributions specifically dealing with the *science system*, the *technology system*, and the *science and technology interface*.

The series of primarily methodological chapters starts with a contribution by ***Edda Leopold***, ***Michael May*** and ***Gerhard Paass***, who provide a general introduction to data and text mining techniques which are useful for analysing large publication and patent databases. Such techniques combine elements from mathematical statistics, machine learning, and information retrieval. The authors present several examples, including one regarding authorship attribution, i.e., classifying documents according to whether they were authored by a specified person or not.

A next contribution deals with issues in patent analysis. Many papers using patent statistics do not accurately define the methodology that was applied. As a consequence different studies on the same topic sometimes produced contradictory results. ***Sybille Hinze*** and ***Ulrich Schmoch*** illustrate how the outcomes of patent analysis depend upon the way in which time scales are defined, the country of origin is identified, patent offices are selected, a patent's quality is measured, and search strategies are conducted. They suggest preferred methodologies and thus contribute to a further standardisation in the field of patent analysis.

The S&T system comprises a wide range of cognitively or technically distinct activities. In order to differentiate and analyse its internal subject heterogeneity analyses of the system should apply adequate subject or content classification systems. Particularly must scientific papers and patents be assigned to scientific disciplines or sub-fields, or grouped into classes on the basis of technical specifications.

Most studies apply existing classification systems, for publications based on a journal category system developed by ISI, and for patents the International Patent Classification (IPC) System. ***Ed Noyons*** illustrates in his contribution how accurate, tailor made subject classifications of documents can be generated, particularly at the level of research topics or specialities. He reviews the potentialities of mapping and data-analytical methods applied in co-word and co-citation analysis, and shows how groupings or clusters obtained can be evaluated in terms of their main actors and the institutions to which they are affiliated.

Collaboration and globalisation are important features of the S&T system. ***Wolfgang Glänzel*** and ***András Schubert*** focus on the science system and show how these phenomena can be studied by analysing co-authorship in scientific publications. They review earlier work on this topic, and present illustrative analyses at the level of individual scientists and that of countries. They depict the global network of science, and discuss empirical, bibliometric findings in terms of the effects that international scientific collaboration may have upon scientific research performance.

A recent development in patent statistics is the renaissance of citation analysis. Citations reflect relations in terms of content and social context and can be used for constructing quality measures. But as derived, indirect indicators they have to be interpreted with caution. ***Bhaven Sampat*** and ***Arvids Ziedonis*** critically examine the motives and functions of patent citations and develop a more differentiated concept of the economic value of patents. Using the example of university patents, they illustrate the different dimensions of value and show a significant relation between patent citations with respect to the probability of licensing.

The participation of women in the S&T system has gained substantial interest and policy relevance during the past decades. ***Fluvio Naldi***, ***Daniela Luzi***, ***Adrianna Valente***, and ***Ilaria Vannini Parenti*** present a methodology that provides a gender classification of authors of scientific publications and inventors of patents, based on their first names. Thus indicators can be calculated of the participation of women in publishing or patenting networks.

Measurement of productivity relates inputs of the S&T system to outputs. Therefore input statistics on spending and human resources are essential elements in a comprehensive system of S&T indicators. ***Marc Luwel*** presents in his contribution a review of the efforts made by the international organisations OECD, UNESCO, and EUROSTAT to generate standardised statistics on R&D input, and discusses major methodological issues. He notes that attempts to calculate per scholarly field productivity measures relating these input measures to output indicators are hampered by the two types statistics giving aggregate measures based on different subject classification systems.

Scientific journals and patents constitute by far the most important data sources in quantitative studies of the S&T system presented in this handbook During the past decade the World Wide Web has become a most important general source of information. More and more scientific and technological information, including publications, are made available and actually retrieved through the web. In addition, bibliometric methods play an important role in the quantitative study of the web, denoted as webometrics.

Therefore the final contribution in the methodology part deals with issues of webometric studies. Peter Ingwersen and Lennart Björneborn wrote it. The authors discuss problems of data collection from the Web, typologies of Web links and numerous conceptual questions. The contribution also briefly addresses Web ‘impact factors’ that bear some resemblance to the well-known journal impact factors published by the *Institute for Scientific Information*.

This handbook's part on the **science system** starts with a contribution by ***Thed van Leeuwen*** on the measurement of academic research performance. He critically discusses academic research assessment exercises carried out in the Netherlands and in the UK, and highlights potentialities and limitations of the use of bibliometric indicators in such assessments. He discusses conditions for proper use of bibliometric indicators in research performance assessments.

Linda Butler investigates changes in the publication behaviour of scientists that the consistent use of bibliometric indicators in the policy domain may induce. She gives a critical view on the effects of such policies for academic output on the example of Australian where a composite index encapsulating a number of performance measures — such as graduate student numbers or completion rates, research income and publication activity — is used to allocate the research component of university block funding.

Michel Zitt and ***Elise Bassecoulard*** present a multi-faceted chapter on internationalisation and globalisation of scientific communication. The authors identify the main engines of science internationalisation, and discuss internationalisation measures applicable to bibliometrics. Internationalisation is studied in terms of journal profiles as well as in the context of scientific collaboration and other networks of interdependencies in science. The authors finally discuss the distribution of knowledge production in the context of internationalisation and the issue of convergence as possible consequence of globalisation.

The increasing mutual dependence amongst science disciplines requiring knowledge flow beyond disciplinary boundaries as well as the fading frontiers between science and technology request increasing attention from all possible perspectives.

The chapter by ***Maria Bordons, Fernanda Morillo*** and ***Isabel Gómez*** provides a bibliometric review of interdisciplinarity in science. Interdisciplinarity is described as the emergence of a new mode of knowledge production, which coexists with the traditional disciplinary science. They analyse, among others, the trend towards interdisciplinarity in scientific research, field-specific characteristics in the context of cross-disciplinary activity, the effect of interdisciplinary research on the relationship among disciplines, and the interaction of interdisciplinarity and bibliometric performance indicators.

Citation analyses are standard methods in bibliometric literature. Conventional bibliometric analyses, however, measure the impact of scientific papers within the research community, in particular, through citations from other scientific papers in the serial literature. In order to provide new indicators of the utility of biomedical research ***Grant Lewison*** studies citations to research papers from different document types, such as clinical guidelines, textbooks, government policy documents, international or national regulations and newspaper articles.

Most studies presented in this handbook relate to the natural sciences, life sciences, and technical sciences. Other domains of human scholarship are discussed by ***Diana Hicks***, who presents a review of methodologies aiming at assessing research performance in social sciences and humanities. Her premise is that bibliometric assessments of research performance in these fields face severe methodological problems. In these fields she identifies four types of literatures, briefly denoted as international journals, books, national journals and the non-scholarly press. She concludes that ignoring the latter three types may produce a distorted picture of social science fields.

During the recent reform of the Chinese scientific system, quantitative evaluation has been introduced into research management and decision making related to national S&T. The number of Chinese publications indexed by the *Science Citation Index* (SCI) has spectacularly increased in the last decade. Nevertheless, most Chinese research results are still published in domestic journals not covered by this database. Therefore, it was decided to develop their own local, Chinese citation indexes. ***Bihui Jin*** and ***Ronald Rousseau*** give a survey of the use of both, the SCI and the Chinese database for the evaluation of national research performance in China.

Macro-indicators, especially national science indicator are standard tools in evaluative bibliometrics. They provide a prompt and comprehensive picture on national research output in science fields under study. But what if trends reveal a decline of national research performance? ***Olle Persson*** and ***Rickard Danell*** show that the breakdown, the *decomposition* of national indicators helps to identify those actors who are most concerned by negative trends, and might support appropriate and targeted policies. They conduct research at different levels of aggregation, particularly at the level of research institutions, research groups, and individual authors. The decomposition method is applied to Swedish neuroscience papers.

The quantitative analysis of trends, performance and structures in technology is a complex task owing to the large variety of technical artefacts and processes. Many studies refer to rather indirect indicators such as foreign trade, labour force, or investment in R&D-intensive sectors. Grupp (1992) suggested, with reference to specific technologies, collecting specification measures and deriving integrated indicators; he labelled this approach as ‘technometrics’. The papers related to technology in this handbook exclusively refer to the use of patent indicators which proved to be a very flexible and powerful analytical tool. Since patent documents were easily accessible through electronic databases, many scholars used patent indicators for different purposes. However, the statistical analysis of patents did not develop to an independent sub-discipline comparable to publication statistics. Hence the term 'patentometrics' suggested by some authors as equivalent to 'bibliometrics' did not become generally accepted.

Patent analyses are not always employed for assessing technology. In particular, in economics the appropriateness and efficiency of the patent system itself is often investigated by means of patent indicators. In this handbook, we focus, rather, on their use for quantitatively analysing technology, and the editors are well aware that some important approaches based on patent indicators are not considered.

Patent indicators provide a favourite tool for analysing the technological performance of countries in a differentiated way. In this context ***Lionel Nesta*** and ***Pari Patel*** suggest a novel indicator combining the analysis of present performance and specialisation with a dynamic perspective. In addition they give a short and comprehensive introduction into the advantages and shortcomings of patent statistics on the country level.

Various scholars have shown on the macro level that in recent decades technology is a major driving force of economy. However, it is quite complex to show the linkage between technological and economic performance on the level of enterprises. In their contribution ***Francis Narin***, ***Anthony Breitzman***, and ***Patrick Thomas*** give convincing evidence that the technological performance of firms measured by patent indicators has a relevant impact on their stock market value. For that purpose they refer to a combined index, which primarily refers to different dimensions of quality rather than pure quantity.

Patent indicators are not only useful at the macro level of countries, but are strategic instruments of firms for assessing the technological orientation and performance of competitors and to benchmark their own competence. In this handbook this type of analysis is represented by two contributions with different approaches. Both chapters illustrate that patents are an important source of strategic information for firms.

Koenraad Debackere and ***Marc Luwel*** present benchmark indicators to assess the technological strengths and weaknesses of companies taking up characteristic elements of economic portfolio analysis. ***Alan Porter*** and ***Nils Newman*** collect a broader set of more straightforward, but informative indicators from patent databases for generating competitive intelligence for technology managers. They highlight the relevance of a careful match of the selected indicators to the specific needs of the users, so that the close interaction of data producers and users is important.

A next contribution addresses novel aspects in the use of patent citations. ***Stefano Breschi*** and ***Francesco Lissoni*** use patent citations for identifying social relations and empirically describing social networks. In their contribution they discuss in detail the methodological appropriateness of this approach and demonstrate its validity by the example of Italian patent applications.

The internalisation of the economy has different aspects such as the increasing foreign trading with technology-intensive goods, production in foreign countries, or the growing R&D activities in foreign countries. As to the latter aspect, the available statistics are sketchy, hardly comparable, and only coarsely differentiated by sectors or fields. ***Dominique Guellec*** and ***Bruno van Pottelsberghe de la Potterie*** impressively show how patent analyses can be used to examine different forms of knowledge flows and thus of the internationalisation of technology.

The last part of the handbook deals with the science and technology interface. ***Elise Bassecoulard*** and ***Michel Zitt*** review various ways of studying this interface. Next, they explore the possibility of relating science and technology on the basis of lexical linkages between articles and patents. It is generally recognized that standard scientific publication and patent subject classification systems do not match. Therefore, the authors particularly examine the possibility of creating correspondence tables between these two types of systems using these lexical linkages.

Robert Tijssen presents a review of the study of the interactions and knowledge flows between science and technology, and focuses on two methodologies. The first is based on citation flows and analyses citations made in patents to the scientific literature and also those from scientific papers to patents. The second can be denoted as a person oriented approach and deals with scientists–inventor relationships, assessing the extent to which authors of scientific publications act as inventors in patents.

Ulrich Schmoch analyses the patent applications of scientific institutions as a proxy for their direct contribution to technology. He shows that these institutions focus their activities on knowledge based fields and that their participation therein is much higher than often assumed.

Stefano Brusoni and ***Aldo Geuna*** approach the S&T interface from an opposite perspective by looking at the science reference of firms in the pharmaceutical sector. For that purpose they analyse the citations in patents to scientific publications. By indicators of specialisation and integration they characterise the different pattern of performance and orientation of the firms analysed. In addition, they apply this multi-dimensional description to characterise the performance of countries.

The discourse on the interaction of science and technology primarily refers to the situation in advanced industrialised countries. In contrast to this general trend, ***Eduardo da Motta e Albuquerque*** examines this topic for less developed countries by using patent and publication indicators. He demonstrates that these indicators are also useful for characterising these countries, and by more detailed investigations for specific countries, he can derive structural indicators supporting the conception of adequate innovation policies.

The current handbook can be considered as the successor of the *Handbook of Quantitative Science and Technology Studies* edited by Anthony van Raan and published in 1988. It is tempting to compare the contents of the two handbooks and to identify major trends in the field during the past 16 years, assuming that both handbooks adequately reflect the state of the art in the field at the time they were published.

A major trend is that publication and patent data have become more widely available for publication analysis and construction of indicators. This reflects developments in information technology during the past two decades. Nowadays large publication and patent databases are available, under certain restrictions, in electronic form. As to publications, the *Science Citation Index* and related Citation Indexes published by the Institute for Scientific Information (currently Thomson Scientific) is the most important database. Many other databases in specific areas are available, most of them, however, without recording citations. As to patents, the major patent offices such as the *US Patent and Trademark Office* or the *European Patent Office* have supported the distribution of patent information through various channels. The launch of electronic information on CD-ROMs stimulated its use for analytical purposes. Many large publication databases and patent databases are currently available through online services, and on CD-ROM, bibliometric macro indicators at the level of countries and scholarly subfields can be purchased as standard indicator products.

In the 1980s there were only three integral ‘bibliometric’ versions of the complete ISI Citations indexes, at ISI and at CHI Research, both located in the U.S., and in Europe at the Hungarian Academy of Sciences. Nowadays several other institutions have such integral versions — or huge extracts from them — that they can use under a number of conditions for large scale bibliometric analysis. This is clearly reflected in the contributions in the current handbook by authors affiliated with these institutions.

At the end of the eighties a broad part of the discourse on indicators focussed on methodological issues, reflecting a latent uncertainty regarding their meaningfulness. Many scholars had no appropriate ideas of how to react to the demand of users for the application of indicators in a policy context. Since that time many experiences have been made of the appropriateness and the practical use of indicators. Today publication and patent indicators tend to be more tailor made, and they are more often designed for answering particular research questions. Indicators appropriate in one research or policy context may be less so in other contexts, and may have to be substituted by other more sophisticated counterparts showing more detail or arranging sub-units from analysed systems in a different way.

If there is any general trend at all in the topics addressed in the various science studies, it is one that reflects a shift from a sociological towards an economic perspective, or from an emphasis on the science system's internal functioning and performance criteria towards an emphasis on the science system's potential technological and economic utility, and on the relationship between science, technology and innovation. Similarly, current technology studies tend to focus more on the technology's science base and on its economic role and value.

Finally, the role of S&T indicators in evaluation and decision making processes in the policy domain has become more prominent. Indicators are not only more frequently produced and more easily available, but are also more frequently used in recent years than they were some 15 years ago. As a consequence several contributions in the handbook propose criteria for their proper use in the sphere of policy, assess the political dimension of S&T indicators, and reflect upon its implications for practitioners in the fields of science and technology studies.

Table 1. Chapters in the Handbook

Attributes; policy context	*Disciplinary approaches; methodologies; types of indicators*	*Authors*
Disciplinary Approaches		
Measuring science	Historical–methodological overview; physical approach	van Raan
Productivity of S&T systems	Econometric, nonparametric approach	Bonaccorsi and Daraio
Use and misuse of S&T indicators for national SD&T policy	Economic policy approach	Grupp and Mogee
Journal gate keeping system	Sociological approach	Braun
S&T policy processes; political dimension of indicators	Socio-political approach	Barré
Technological paradigms and long term trajectories	Economic–evolutionary approach using patent data	Andersen

Attributes; policy context	*Disciplinary approaches; methodologies; types of indicators*	*Authors*
Access to S&T information for developing countries	Information–scientific approach with emphasis on Open Access	Arunachalam
General Methodology		
Data and text mining	Statistical, machine learning, and information retrieval approaches	Leopold, May and Paass
Basic characteristics of patent data	Patent time scale, country of origin, office, quality and search strategies	Hinze and Schmoch
Socio-cognitive structures in S&T activities	Mapping techniques; co-word, co-citation analysis	Noyons
Scientific collaboration; globalisation	Co-authorship links; multinational research articles	Glänzel and Schubert
Economic value of patents	Patent citation analysis	Sampat and Ziedonis
Participation of women in S&T	Use of first names of authors or inventors for gender classification	Naldi, Luzi, Valente and Vannini Parenti
R&D input data	Efforts by OECD; combination with macro bibliometric indicators	Luwel
Studying the World Wide Web	Data collection; link typologies; conceptual issues; impact factors	Ingwersen and Björneborn
The Science System		
Academic research performance	Conditions for proper use of indicators in research assessments	van Leeuwen
Effects of funding formula upon scientists' publication practices	Use of publication counts in Australian academic funding	Butler
Internationality of science	Journal internationalisation indexes; international co-authorships; disciplinary specialisation profiles	Zitt and Bassecoulard
Multi- and inter-disciplinarity of research	Co-classification of journals; cross-disciplinary co-authorships and citations	Bordons, Morillo and Gómez

Attributes; policy context	*Disciplinary approaches; methodologies; types of indicators*	*Authors*
Practical effects of biomedical research	Citations from clinical guidelines, textbooks, regulations and newspapers	Lewison
Performance in social sciences and humanities	Publication analyses of four types of literatures	Hicks
Chinese science system	Publication based indicators from ISI and Chinese Citation databases	Jin and Rousseau
Scandinavian science system	Breakdown of macro indicators in terms of institutions and authors	Persson and Danell
The Technology System		
Performance of national innovation systems	Patent performance and dynamics by country and by technology sector	Nesta and Patel
Technological and stock market performance	Patent and stock market statistics by company	Narin, Breitzman, Thomas
S&T portfolio management of companies and regions	Patent statistics measuring relative technological specialisation	Debackere and Luwel
Competitive technical intelligence for company managers	Manipulating information from patent databases	Porter and Newman
Knowledge networks in technological innovation	Patent citations; co-inventions, social network analysis	Breschi and Lissoni
Internationalisation of technology	Foreign inventors of domestic applicants and v.v.; international co-inventorship	Guellec and Pottelsberghe de la Potterie
The Science–Technology Interface		
Correspondence tables between patent and scientific classifications	Lexical linkages between research articles and patents	Bassecoulard and Zitt
Knowledge flows between science and technology	Scientist–inventor relationships; citations from articles to patents	Tijssen
Contribution of public non-profit science institutions to technology	Scientists as inventors of patents in science intensive fields	Schmoch

Attributes; policy context	*Disciplinary approaches; methodologies; types of indicators*	*Authors*
Knowledge specialisation and integration of companies and countries	Publication, citation and patent based indicators of depth and breadth of a knowledge base	Brusoni and Geuna
S&T systems in developing countries	Differentiation of countries using statistics based on patents and research papers	da Motta e Albuquerque

REFERENCES

Garfield, E (1979). *Citation Indexing*. New York: Wiley.

Grupp, H. (1992). The measurement of technical performance of innovations by technometrics and its impact on established technology indicators. *Research Policy*, 23, 175–193.

Van Raan, A.F.J (1988) (ed). *Handbook of Quantitative Science and Technology Studies*. Amsterdam: North Holland.

Chapter 1

MEASURING SCIENCE

Capita Selecta of Current Main Issues

Anthony F.J. van Raan

Centre for Science and Technology Studies (CWTS), Leiden University, the Netherlands
E-mail: vanraan@cwts.leidenuniv.nl

Abstract: After a review of developments in the quantitative study of science, particularly since the early 1970s, I focus on two current main lines of 'measuring science' based on bibliometric analysis. With the developments in the Leiden group as an example of daily practice, the measurement of research performance and, particularly, the importance of indicator standardisation are discussed, including aspects such as interdisciplinary relations, collaboration, 'knowledge users'. Several important problems are addressed: language bias; timeliness; comparability of different research systems; statistical issues; and the 'theory–invariance' of indicators. Next, an introduction to the mapping of scientific fields is presented. Here basic concepts and issues of practical application of these 'science maps' are addressed. This contribution is concluded with general observations on current and near-future developments, including network–based approaches, necessary 'next steps' are formulated, and an answer is given to the question 'Can science be measured?'

1. TOWARD A METRIC OF SCIENCE REVISITED[1]

From the early sixties onwards we see a strong increase in quantitative material on the state-of-the art in science and technology. National institutes of statistics, UNESCO, OECD, and the European Commission are main examples of organisations starting to collect systematically data on the

[1] The book *Toward a metric of Science: The Advent of Science Indicators* (Elkana et al., 1978) has always been a one of my major sources of inspiration. This contribution to the Handbook is based on earlier publications by the author (Van Raan 2000a; Van Raan and Noyons 2002).

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 19-50.

development of science and technology. An important milestone is the first issue of the OECD 'Frascati Manual' (OECD, 1963), a handbook devoted to the development of a standard practice for surveys of the measurement of scientific and technical activities. At the same time, and strongly related to this data explosion, the quantitative appraisal of current science gains influence. As a genre in the study of the history of science, the quantitative approach of the development of science, 'scientometrics', is certainly not new. A remarkable early piece of work is "Histoire des sciences et des savants depuis deux siècles". The author, Alphonse de Candolle (1873), described the changes in the scientific strength of nations by membership of scientific societies, and he tried to find 'environmental factors' of all kinds (even including the role of the celibate) for the scientific success of a nation. Later, in the 1920s, Lotka (1929) published his famous work on the productivity of chemistry researchers. Here scientometrics is clearly differentiated into 'bibliometrics'.

Undoubtedly the invention of the *Science Citation Index* by Eugene Garfield is a major breakthrough (Wouters, 1999). This invention enabled statistical analyses of the scientific literature on a very large scale. It marks the rise of bibliometrics as a powerful field within the studies of science. Such great scientists as Derek de Solla Price and Robert Merton recognised the value of Garfield's invention, Price from the perspective of contemporaneous history of science, Merton from the perspective of normative sociology.

Scientists are fascinated by basic features such as simplicity, symmetry, harmony, and order. The *Science Citation Index* enabled De Solla Price to start with the development of a 'physical approach' to science, in which he tried to find laws to predict further developments, inspired by the ideas of Newtonian and statistical mechanics. In this perspective, quantitative measures of science, '*indicators',* are guides to find and, as a crucial next step, to understand such basic features. The most basic feature concerns the cognitive dimension: the development of content and structure of science. More on the mundane surface science indicators relate to the social dimension of science, in particular to aspects formulated in questions such as 'How many researchers? How much money is spent on science? How 'good' are research groups? How does communication in science work, particularly what is the role of books, journals, conferences (Borgman, 1990)? And longer than we often realise there is another question: 'What is the economic profit of scientific activities?' A landmark in the development of science indicators is the first publication in a biennial series of the *Science Indicators* Report in 1973. Stimulated by the success achieved by economists in developing quantitative measures of political significance (e.g., unemployment, GNP), the US National Science Board started this indicator report series in which we find more emphasis on the demographic and economic state of science than on the cognitive state of science (National Science Board, 1973).

Making quantitative indicators of anything thinkable fascinates some people and horrifies others as being nonsense and taking us back to the cabbalistic magic number world of Paracelsus. But there are famous classical pronouncements to support the attempt to measure things. Horace (65–5 BC): “There is a measure in all things” (Est modus in rebus), Johannes Kepler (1597): “The mind comprehends a thing the more correctly the closer the thing approaches toward pure quantity as its origin”, and, from the place where I live and work, Leiden, the discoverer of superconductivity, Heike Kamerlingh Onnes (1882): “Measuring is knowing”.

There is no final theory of science providing *the* methodology of measurement. It is a returning hype in the social studies of science to incite the scientific community with this observation. But are we really troubled by this poverty of theoretical content? I don’t think so (van Raan, 1997). Do not expect a classical mechanics of scientometrics. With very high probability: it does not exist. The absence of any explicit theory to guide the making and use of indicators may not be good, but the adoption of a single one, for instance, a trendy dominating ‘theory’, is likely to be worse (Holton, 1978). It is normal practice in empirical science to begin a search without a theoretical clarification and try to establish a model to explain the findings later. Certainly in such measurements we do have at least implicit basic ideas about ‘how things work’ and the same is true for the construction and use of science indicators. Therefore it is crucial to make these implicit assumptions clear to the outside world. This will allow us to turn the absence of a general theory of the development of science into a very profitable situation, in the words of Gerald Holton: ‘perhaps indicators may be developed eventually that are *invariant* with respect to theoretical models. They and only they allow rival theories to be put to empirical tests’. To put it more bluntly: we cannot develop a sound theoretical model of the ‘sociology of knowledge’ yet, as we simply need more empirical work based on the richness of available and future data in order to develop a better quantitative understanding of the processes by which science and society mutually influence each other’s progress. In this contribution I will argue that advanced bibliometric indicators approach the above characteristic of invariance.

What is the difference between data and indicators? An indicator is the result of a specific mathematical operation (often simple arithmetic) with data. The mere number of citations of one publication in a certain time period is *data*. The measure in which such citation counts of all publications of a research group in a particular field are normalised to citation counts of all publications worldwide in the same field, is an *indicator*. An indicator is a measure that explicitly addresses some assumption. In our example the assumption is: this is the way to calculate the international scientific influence of a research group. So, to begin with, we need to answer the question: what features of science can be given a numerical expression? Thus indicators can not exist without a specific goal in mind, they have to address specific questions, and thus they have to be created to gauge important 'forces'; for example, how scientific progress is related to specific cognitive as well as socio-economic aspects. Indicators must be problem driven, otherwise they are useless. They have to describe the recent past in such a way that they can guide us, can inform us about the near future. A second and more fundamental role of indicators is their possibility to test aspects of theories and models of scientific development and its interaction with society. In this sense, indicators are not only tools for science policy makers and research managers, but also instruments in the study of science. But we also have to realise that science indicators do not answer typical epistemological questions such as: How do scientists decide what will be called a scientific fact? How do scientists decide whether a particular observation supports or contradicts a theory? How do scientists come to accept certain methods or scientific instruments as valid means of attaining knowledge? How does knowledge selectively accumulate? (Cole et al., 1978).

De Solla Price (1978) strikingly described the mission of the indicator maker: find the most simple pattern in the data at hand, and then look for the more complex patterns which modify the first. What should be constructed from the data is not a number but a pattern, a cluster of points on a map, a peak on a graph, a correlation of significant elements on a matrix, a qualitative similarity between two histograms. If these patterns are found the next step is to suggest models that produce such patterns and to test these models by further data. A numerical indicator or an indicative pattern, standing alone, has little significance. The data must be given perspective: the change of an indicator with time, or different rates of change of two different indicators. Crucial is that numerical quantities are replaced by geometrical or topological objects or relations (Ziman, 1978).

We know already from the early indicator work that these 'simple patterns' exist: the rank of countries by the number of publications is remarkably stable from year to year (Braun et al., 1995). The absolute size of the scientific research activity in the number of publications of any nation is in very good agreement with its electrical power consumption in kilowatt-hours, indicating that scientific power, economic power, and national wealth are strongly related.

More or less at the same time as the above thoughts on the metric of science, Francis Narin coined the concept of ‘evaluative bibliometrics’. His pioneering work on the development of research performance indicators (Narin, 1976, 1978), mainly on the macro level, i.e., the performance of countries, was a new, important breakthrough which contributed substantially to the measurement of scientific activities. In 1978 Tibor Braun founded the journal *Scientometrics*. This event marks the emancipation of the field of quantitative studies of science. Also in journals such as *Research Policy* and the *Journal of the American Society for Information Science* we find more and more publications about ‘measuring science’, and most of them are on topics that are still very relevant. We mention, without being exhaustive, the seminal papers in the 1970s on the development of ‘relational’ methods such as co-citation analysis for the mapping of scientific fields (Small, 1973), on scientific collaboration by deB. Beaver and colleagues (Beaver, 1978), on measuring the growth of science (Moravcsik, 1975; Gilbert, 1978), the meaning of citation patterns for assessing scientific progress (Moravcsik and Murugesan, 1978), and on mobility in science (Vláchy, 1979).

In the early eighties we see the rapid rise of co-citation analysis (Small and Greenlee, 1980; Sullivan et al., 1980; Price, 1981; White and Griffith, 1981; Noma, 1982; McCain, 1984) and of co-word analysis (Callon et al., 1983; Rip and Courtial, 1984), an increasing emphasis on advanced statistical analysis of scientometric parameters (Haitun, 1982; Schubert and Glänzel, 1983), the application of bibliometric methods in the social sciences (Peritz, 1983), indicators of interdisciplinary research (Porter and Chubin, 1985), and comparison of peer opinions and bibliometric indicators (Koenig, 1983).

An important further breakthrough was the work of Martin and Irvine (1983) on the application of science indicators at the level of research groups. Around the same time (the beginning of the eighties) our Leiden institute had also started with bibliometric analysis oriented on research groups (Moed et al., 1983) and Braun and co-workers focused on the scientific strength of countries in a wide range of research fields (Braun et al., 1988).

Now, almost thirty years after Narin’s *Evaluative Bibliometrics*, twenty-five years after the publication of *Toward a Metric of Science: The Advent of Science Indicators* (Elkana et al., 1978), twenty years after Martin and Irvine (1983), and fifteen years after the *Handbook of Quantitative Studies of Science and Technology* (van Raan, 1988) we may state *plus ça change, plus c’est la même chose*. What changed is the very significant progress in application oriented indicator work based on the enormous increase of available data and, above all, the almost unbelievable, compared with the situation in the seventies, increase of computing power and electronic facilities. I hope this contribution and handbook as a whole will prove this progress convincingly.

What also changed is the method of publishing. Electronic publishing and electronic archives mark an area of new information technology. I expect that most changes will be primarily technological but not conceptual. Publication via journals of high reputation is in most fields of science crucial for receiving professional recognition. That will remain so in the rapidly developing electronic area. A much more revolutionary change in science is the increasing availability and sharing of research results and, particularly, research data.

What remained, however, are some of the most fundamental questions. For instance: do science maps derived from citation and/or concept–similarity data have reality in a strictly spatial sense? In other words, do measures of similarity imply the existence of a metric space? This question brings us to an even more fundamental problem: the ontological status of maps of science will remain speculative until more has been learned about the structure of the brain itself (de Solla Price, 1978). For instance, it remains fascinating that science can be represented quite well in 2D space. Why is that so? Because our own brain is a (folded) two dimensional structure?

And yes, some old wishes have come true. It is now possible to make a time series of science maps, a 'science cinematography' that enables us to examine shifts in clusters over time and to investigate the nature of change of research themes and specialties. Short term extrapolation may be feasible.

A new development is a 'physical' network approach to analysing publication and citation relations. Recently we reported some first results on network characteristics of a reference based, bibliographically coupled publication network structure (van Raan, 2003). It was found that this network of clustered publications shows different topologies depending on the age of the references used for building the network. Also progress is made in the understanding of the statistics of citation distributions. This is of crucial importance, as it is directly related to the 'wiring' (citations) of the 'nodes' (publications) in the network structure of science. A two-step competition process is applied as a model for explaining the distribution of citations ('income') over publications ('work'). A distribution function of citing publications is found which corresponds very well to the empirical data. It is not a power law, but a modified Bessel function. This model has a more generic value, particularly in economics for explaining observed income distributions (van Raan, 2001).

In this contribution we focus on two main lines of 'measuring science' based on bibliometric analysis. First, in the next section, we discuss the measurement of research performance, including aspects such as interdisciplinarity, collaboration, 'knowledge users'. I address several important problems: language bias; timeliness; comparability of different research system; statistical issues; the relation between bibliometric finding and peer judgements. The latter issue is followed by a first discussion of Holton's ideal of 'theory invariant' indicators. In Section 3 an introduction to the mapping of scientific fields is presented. I discuss basic concepts and issues of the practical application of these 'science maps'. Finally, in Section 4 this contribution is concluded with some general observations on current and near-future developments, particularly in relation to network–based approaches and growth phenomena. Necessary 'next steps' are formulated. But first, back to the basics.

2. BIBLIOMETRIC MEASUREMENT OF SCIENTIFIC PERFORMANCE

2.1 Basic Concepts

The rationale of the bibliometric approach to measuring scientific performance presented in this contribution is as follows. Scientific progress can be defined as the substantial increase of our knowledge about 'everything'. In broad outline we discern basic knowledge ('understanding') and applicable knowledge ('use'). This knowledge can be tacit ('craftsmanship') or codified ('archived & publicly accessible'). Scientists have communicated (and codified) their findings in a relatively orderly, well defined way since the 17th century. Particularly is the phenomenon of serial literature crucial: publications in international journals. Thus communication, i.e., exchange of research results, is a crucial aspect of the scientific endeavour. Publications are not the only, but certainly very important elements, in this process of knowledge exchange.

Each year about 1,000,000 publications are added to the scientific archive of this planet. This number and also numbers for sub-sets of science (fields, institutes) are in many cases sufficiently high to allow quantitative analyses yielding statistically significant findings. Publications offer usable elements for 'measuring' important aspects of science: author names, institutional addresses, journal (which indicates not only the field of research but also status!), references (citations), concepts (keywords, keyword combinations). Although not perfect, we adopt a publication as a 'building block' of science and as a source of data. This approach clearly defines the basic assumptions of bibliometrics (Kostoff, 1995). Thus bibliometric assessment of research performance is based on one central assumption: scientists who have to say something important do publish their findings vigorously in the open international journal ('serial') literature. This choice introduces unavoidably a 'bibliometrically limited view of a complex reality'. For instance, journal articles are not in all fields the main carrier of scientific knowledge; they are not 'equivalent' elements in the scientific process, they differ widely in importance; and they are challenged as the 'gold standard' by new types of publication behaviour, particularly electronic publishing. However, the daily practice of scientific research shows that inspired scientists in most cases, and particularly in the natural sciences and medical research fields, go for publication in the better and, if possible, the best journals. A similar situation is developing in the social and behavioural sciences (Glänzel, 1996; Hicks, 1999), engineering and, to a lesser extent, in the humanities. This observation is confirmed by many years of experience in peer review based research evaluation procedures.

Work of at least some importance provokes reactions of colleagues. They are the international forum, the 'invisible college', by which research results are discussed. Often these colleagues play their role as a member of the invisible college by referring in their own work to earlier work of other scientists. This process of citation is a complex one, and it certainly does not provide an 'ideal' monitor on scientific performance (MacRoberts and MacRoberts, 1996). This is particularly the case at a statistically low aggregation level, e.g., the individual researcher. But the application of citation analysis to the work, the 'oeuvre' of *a group of researchers as a whole over a longer period of time*, does yield in many situations a strong indicator of scientific performance.

Citation analysis is based on reference practices of scientists. The motives for giving (or not giving) a reference to a particular article may vary considerably (Brooks, 1986; MacRoberts and MacRoberts, 1988; Vinkler, 1998). There is, however, sufficient evidence that these 'reference motives' are not so different or 'randomly given' to such an extent that the phenomenon of citation would lose its role as a reliable measure of impact (van Raan, 1998).

Why bibliometric analysis of research performance? Peer review undoubtedly is and has to remain the principal procedure of quality judgment. But peer review and related expert–based judgments may have serious shortcomings and disadvantages (Moxham and Anderson, 1992; Horrobin, 1990). Subjectivity, i.e., dependence of the outcomes on the choice of individual committee members, is one of the major problems. This dependence may result in conflicts of interests, unawareness of quality, or a negative bias against younger people or newcomers to the field. Basically, the methodological problem of determining the quality of a subject is still far from solved, as illustrated by the results of re-review of previously granted research proposals, see, for instance, Nederhof (1988). I do not plead for a replacement of peer review by bibliometric analysis. Subjective aspects are not merely negative. In any judgment there must be room for the intuitive insights of experts. I claim, however, that for a substantial improvement of decision making an advanced bibliometric method, such as presented in this contribution has to be used in parallel with a peer–based evaluation procedure.

The earlier mentioned pioneering work of Narin (1976) and of Martin and Irvin (1983) clearly showed that the most crucial parameter in the assessment of research performance is international scientific influence. Citation–based bibliometric analysis provides indicators of international impact, influence. This can be regarded as, at least, one crucial aspect of scientific quality, and thus a 'proxy' of quality as follows from a long standing experience in bibliometric analysis. Perhaps this is the best answer of the classical question posed by Eugene Garfield (1979): 'Is citation analysis a legitimate evaluation tool?' Therefore we have developed standardised bibliometric procedures for assessing research performance within the framework of international influence. Undoubtedly, this approach does not provide us an ideal instrument, working perfectly in all fields under all circumstances. But the approach presented in this contribution works very well in the large majority of the natural, the medical, the applied, and the behavioural sciences. These fields of science are the most cost intensive and the ones with the strongest socio-economic impact. For a recent application of bibliometric research performance assessment in a typical applied field such as food and nutrition research we refer to Van Raan and Van Leeuwen (2002). The application of bibliometric analysis in the humanities is discussed by Moed et al. (2002).

A first and good indication of whether bibliometric analysis is applicable to a specific field is provided by the publication characteristics of the field; in particular, the role of *international* refereed journals. If international journals are a dominating or at least a major means of communication in a field, then in most cases bibliometric analysis is applicable. Therefore it is important to study the 'publication practices' of a research group, department, or institute, in order to establish whether bibliometric analysis can be applied. A practical measure here is the share of CI-covered[2]

[2] The Science Citation Index, the Social Science Citation Index, the Arts & Humanities Citation Index, and the 'specialty' citation indexes (CompuMath, Biochemistry and Biophysics, Biotechnology, Chemistry, Material Science, Neurosciences) are produced and published by the Institute for Scientific Information (ISI/Thomson Scientific) in

publications in the total research output. For 'not-CI covered publications' a restricted type of analysis is possible, in so far as these publications are cited by articles in journals covered by the CI.

We have already noticed that journal publications are challenged as the 'gold standard' in science as the worldwide web has changed scientific communication. Researchers use the web for information seeking, and in addition to the above mentioned 'not-CI covered publications' there is an enormous number of further publications and data included in institutional and personal websites. Thus next to citation analysis, in the use of data provided via the internet, 'webometrics' offers interesting additional opportunities to aid citation–based bibliometric analysis in evaluation and mapping approaches (Björneborn and Ingwersen, 2001; Bar-Ilan, 2001; Thelwall and Smith, 2002; Thelwall and Harries, 2003).

The Leiden group has gained an extensive experience in bibliometric analysis. In a period of almost 20 years we have studied the research performance of many thousands of research groups, worldwide. By all these activities an empirical gold mine was created. We first discuss our methodology in the next section, and in Section 2.3.5 we explain why we think that this methodology has yielded indicators which, at least, approach Holton's ideal of theory–invariant measures.

2.2 Details of the Methodology

One of the most crucial objectives in bibliometric analysis is to arrive at a consistent and *standardised set* of indicators. The methodology presented in this section is driven by this motive. Research output is defined as the number of articles of the institute, as far as covered by the *Science Citation Index* (SCI) and all its related databases (see footnote 3). As 'article' we consider the following publication types: normal articles (including proceedings papers published in journals); letters; notes; and reviews (but not meeting abstracts, obituaries, corrections, editorials, etc.).

I take the results of a recent analysis by our institute of a German medical research institute as an example (over the period 1992–2000). Table 1.1 shows the number of papers published, ***P***, which is also a first indication of the size of an institute. This number is about 250 per year. Next we find the total number of citations, ***C***, received by ***P*** in the indicated period, *and corrected for self-citations*. For papers published in 1996 citations are counted during the period 1996–2000, for 1997 papers citations in 1997–

Philadelphia. Throughout this paper we use the term 'CI' (Citation Index) for the above set of databases.

2000, and so on. For the outsider this looks like 'just counting numbers'. But the reliable establishment of even these two basic indicators is far from trivial. Verification is crucial in order to remove errors and to detect incompleteness of addresses of research organisations, departments, groups.

In citation analysis an entire range of pitfalls and sources of error is lurking. We refer to Van Raan (1996) for the many methodological and technical problems which have to be solved in order to conduct a bibliometric analysis properly. There is ample empirical evidence that in the natural and life sciences, basic as well as applied, the average 'peak' in the number of citations is in the third or fourth year after publication. Therefore a five-year period is appropriate for impact assessment. A trend analysis is then based on 'moving' and partially overlapping five-year periods, as presented in Table 1.1.

The third and fourth indicators are the average number of citations per publication ***(CPP)***, again without self-citations, and the percentage of not-cited papers, ***% Pnc***. We stress that this percentage of non-cited papers concerns, like all other indicators, the given time period. It is possible that publications not cited within such a time period will be cited after a longer time. This is clearly visible when comparing this indicator for the five-year periods (e.g., 1996–2000: 30%) with that of the whole period (1992–2000: 21%). The values found for this medical research institute are quite normal.

How do we know that a certain number of citations, or a certain value of citations-per-publication is low or high? To answer this question we have to make a comparison with (or normalisation to) a well chosen international reference value, and thus to establish a reliable measure of *relative, internationally field–normalised impact*. Another reason for normalising the measured impact of an institute (***CPP***) to international reference values is that overall worldwide citation rates are increasing. I stress, however, that the distribution of citations over publications is skew and therefore we have to be careful with the use of mean values. In Section 2.3 a short discussion of statistical problems in bibliometric analysis is given.

First, the average citation rate of all papers (worldwide) in the journals in which the institute has published (***JCSm***, the **m**ean **J**ournal **C**itation **S**core of the institute's 'journal set', and ***JCS*** for one specific journal) is calculated. Thus this indicator ***JCSm*** defines a worldwide reference level for the citation rate of the institute. It is calculated in the same way as ***CPP***, but now for all publications in a set of journals (see van Raan, 1996, 2003). A novel and unique aspect is that we take into account the type of paper (e.g., letters, normal article, review) *as well as* the specific years in which the papers were published. This is necessary, because the average impact of journals may have considerable annual fluctuations and large differences per article type, see Moed and Van Leeuwen (1995, 1996).

With help of the ratio ***CPP/JCSm*** we observe whether the measured impact is *above* or *below* the international average. However, comparison of the institute's citation rate ***(CPP)*** with the average citation rate of its journal set (***JCSm***) introduces a specific problem related to journal status (Lewison, 2002). For instance, if a research group publishes in prestigious (high impact) journals, and another group in rather mediocre journals, the citation rate of articles published by both groups may be equal *relative to* the average citation rate of their respective journal sets. But generally one would argue that the first group evidently performs better than the second. Therefore we developed a second international reference level, a *field*–based world average ***FCS***, and ***FCSm*** in the case in which more fields are involved. This indicator is based on the citation rate of *all* papers (worldwide) published in *all* journals of the field(s)[3] in which the institute is active, and not only the journals in which the institute's researchers publish their papers. Thus, for a publication in a less prestigious journal one may have a (relatively) high ***CPP/JCSm*** but a lower ***CPP/FCSm*****,** and for a publication in a more prestigious journal one may expect a higher ***CPP/FCSm*** because publications in a prestigious journal will generally have an impact above the field–specific average.

Table 1.1. Bibliometric analysis of a medical research institute, 1992–2000

Period	*P*	*C*	*CPP*	*%P nc*	*CPP/ JCSm*	*CPP/ FCSm*	*CPP/ D-FCSm*	*JCSm/ FCSm*	*% sc*
1992 – 00	2,245	43,665	19.45	21	1.26	1.95	1.85	1.55	18
1992 – 96	1,080	11,151	10.33	36	1.27	2.02	1.95	1.58	22
1993 – 97	1,198	12,794	10.68	34	1.24	2.03	1.92	1.63	21
1994 – 98	1,261	12,217	9.69	32	1.19	1.85	1.72	1.55	22
1995 – 99	1,350	13,709	10.15	31	1.21	1.89	1.76	1.56	21
1996 – 00	1,410	14,815	10.51	30	1.20	1.91	1.76	1.59	21

[3] We use here the definition of fields based on a classification of scientific journals into *categories* developed by ISI. Although this classification is not perfect, it provides a clear and 'fixed' consistent field definition suitable for automated procedures within our data system. A more 'real world', user oriented, definition of fields can be provided by the bibliometric mapping methodology discussed in Section 3 of this contribution.

The same procedure is used as applied in the calculation of ***JCSm***. Often an institute is active in more than one field. In such cases a weighted average value is calculated, the weights being determined by the total number of papers published by the institute in each field. For instance, if the institute publishes in journals belonging to genetics as well as to cell biology, then the ***FCSm*** of this institute will be based on both field averages. Thus the indicator ***FCSm*** represents a *world average*[4] in a specific (combination of) field(s). It is also possible to calculate ***FCSm*** for a specific country or for the European Union. The example discussed in this paper concerns a German medical research institute, and for this institute we calculated the Germany–specific ***FCSm*** value, ***D-FCSm***.

As in the case of ***CPP/JCSm***, if the ratio ***CPP/FCSm*** is above 1.0 the impact of the institute's papers exceeds the field–based (i.e., *all* journals in the field) world average. We observe in Table 1.1 that the ***CPP/JCSm*** is 1.20, ***CPP/FCSm*** 1.91 and ***CPP/D-FCSm*** is 1.76 in the last period 1996–2000. These results show that the institute is performing well above international average. The ratio ***JCSm/FCSm*** is also an interesting indicator. If it is above 1.0, the mean citation score of the institute's journal set exceeds the mean citation score of all papers published in the field(s) to which the journals belong. For the institute this ratio is around 1.59. This means that the institute publishes in journals with, generally, a high impact. The last indicator shows the percentages of self-citations (***%Sc***). About thirty percent is normal, so the self-citation rates for this institute are certainly not high (about 20%).

A general, and important, observation is the 'stability' over time of most indicators. This is quite typical, particularly for groups and institutes of high reputation. The conclusion to be drawn from this observation is that the indicators are not a 'noisy set of measures' but apparently represent an enduring characteristic of scientific work, including communication practices.

I regard the *internationally standardised impact indicator* ***CPP/FCSm*** as our 'crown' indicator. This indicator enables us to observe immediately whether the performance of a research group or institute is significantly far below (indicator value < 0.5), below (indicator value between 0.5 and 0.8), about (between 0.8 and 1.2), above (between 1.2 and 1.5), or far above (>1.5) the international impact standard of the field. I stress, however, that

[4] About 80 percent of all CI-covered papers is authored by scientists from the United States, Western Europe, Japan, Canada, and Australia. Therefore our 'world average' is dominated by the Western world.

for the interpretation of the measured impact value one has to take into account the *aggregation level of the entity* under study. The higher the aggregation level the larger the volume of publications and the more difficult it is to have an impact significantly above the international level. Based on our long standing experiences, I can say the following. At the 'meso level' (e.g., a university, faculty, or large institute, with about 500 or more publications per year), a ***CPP/FCSm*** value above 1.2 means that the institute's impact as a whole is significantly above the (western) world average. With a ***CPP/FCSm*** value above 1.5, such as in our example, the institute can be considered to be scientifically strong, with a high probability of finding very good to excellent groups. Thus the next step in a research performance analysis is a breakdown of the institution into smaller units, i.e., research groups. Therefore the bibliometric analysis has to be applied on the basis of institutional input data about personnel and composition of groups. The algorithms then can be repeated on the lowest but most important aggregation level, the research group. In most cases the volume of publications at this level is 10 to 20 per year.

Particularly at this lower aggregation level the verification of the data is crucial (e.g., correct assignment of publications to research groups, completeness of publications sets). In our institute we have developed standardised procedures for carrying out the analysis as conscientiously as possible. These procedures are discussed thoroughly beforehand with the client institutes.

At the group level a ***CPP/FCSm*** value above 2 indicates a very strong group, and above 3 the groups can be, generally, considered to be excellent and comparable to the top groups at the best US universities. If the threshold value for the ***CPP/FCSm*** indicator is set at 3.0, excellent groups can be identified with high probability (van Raan, 2000a). As an additional indicator of *scientific excellence* the number of publications within the *top 10%* of the worldwide impact distribution of the field concerned is determined for the target entity (see Noyons et al., 2003). In the calculation of this indicator the entire citation distribution function is taken into account, thus providing a better statistical measure than those based on mean values (see Section 2.3).

Science is, for a major part, teamwork. Particularly is international collaboration essential, not only for the working floor but also as policy for countries to keep pace in scientific progress (Vinkler, 1993; Arunachalam et al., 1994; Melin and Persson, 1996; Glänzel, 2001). For all the above indicators we also perform a breakdown into types of *scientific co-operation* according to the publication addresses: work by only the unit itself; in a national collaboration; or in an international collaboration. Generally one observes the highest impact for publications in international collaboration.

A further important step is the *breakdown* of the institute's output *into* research fields. This provides a clear impression of the research scope or 'profile' of the institute. Such a *spectral analysis* of the output is based on the simple fact of an the institute's researchers publishing in journals of many different fields. Our example, the German medical research institute, is a centre for molecular research oriented towards medicine. The researchers of this institute are working in a typical interdisciplinary environment. The institute's publications are published in a wide range of fields: biochemistry and molecular biology, genetics and heredity, oncology, cell biology, and so on. By ranking fields according to their size (in terms of numbers of publications) in a graphical display, we construct the research profile of the institute. Furthermore, we provide the field–normalised impact values of the institute's research in these different fields with help of ***CPP/FCSm***.

Figure 1.1 shows the results of this *bibliometric spectroscopy*. Thus it becomes immediately visible in which fields within its interdisciplinary research profile the institute has a high (or lower) performance. We observe the scientific strength of the target institute: its performance in the top four fields is high to very high. If we find a smaller field with a relatively low impact (i.e., a field in the lower part, the 'tail' of the profile), this does not necessarily mean that the (few) publications of the institute in this particular field are 'bad'. Often these small fields in a profile are those that are quite 'remote' from the institute's core fields. They are, so to say, peripheral fields. In such a case the institute's researchers may not belong to the dominating international research community of those fields, and their work may be not be cited as frequently as the work of these dominating ('card holding') community members.

In a similar way a breakdown of *the citing publications* into fields of science is made, which yields a profile of the *users* of scientific results (as far as represented by citing publications). This 'knowledge users' profile is a powerful indicator of *who* is using *which* research results, *where* (in which fields) and *when*. Thus it analyses *knowledge diffusion and knowledge use* and it indicates further interdisciplinary 'bridges', potential collaboration, and possible 'markets' in the case of applied research. For an example of these 'knowledge user profiles' I refer to Van Raan and Van Leeuwen (2002). The construction of these profiles can be considered also as an empirical method of studying interdisciplinary aspects of research. For instance, the distribution of the lengths of the field–specific bars in the profile can be used as a measure of interdisciplinarity.

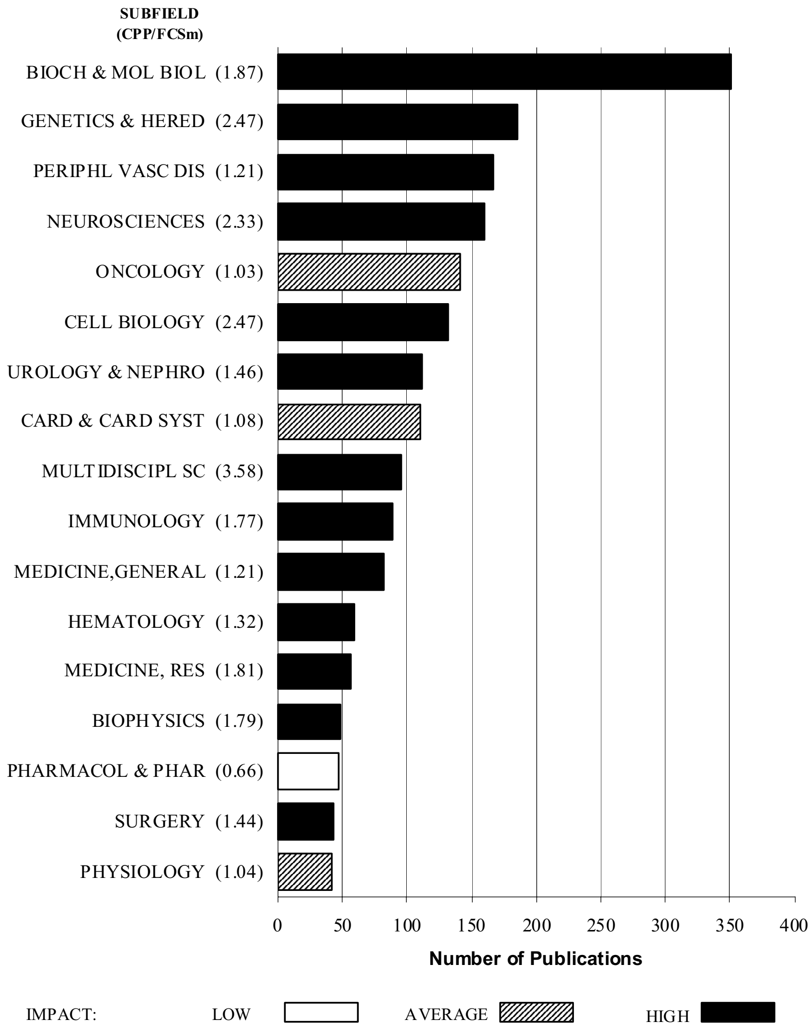


Figure 1.1. Research profile of a medical research institute, 1992–2000

2.3 Important Issues in Applications…and What About Theory?

2.3.1 Language bias

Recent work (Grupp et al., 2001; van Leeuwen et al., 2001) shows that the utmost care must be taken in interpreting bibliometric data in a comparative evaluation of national research systems (May, 1997). The measured value of impact indicators of research activities at the level of an institution and even of a country strongly depends upon whether one includes or excludes publications in CI-covered journals written in languages other than English. This is owed to the simple fact of the CI covering non-English language journals of which the papers have a considerably lower impact than those in the English language journals. Differences of measured impact of the order of 10 to 20% are possible. These findings clearly illustrate that indicators, even at the 'macro level', need to be interpreted against the background of their inherent limitations, such as, in this case, the effects of the language of publication.

2.3.2 Timeliness of the analysis

A frequently posed question concerns the 'delay problem': Does bibliometric analysis suffer from a substantial 'delay' in the measurement of research performance (Egghe and Rousseau, 2000)? An answer to this question first needs a further refinement: delay compared to what? To the average 'processing time' of a publication? To the average 'running time' of a project? Or to peer review 'time cycles'? The entire process starting with scientific activities and leading to 'publishable' results, the writing of an article, the submission of the article, the publication of the article, the citations to the article, varies considerably for the different fields of science, and often within a field. Depending on type of activities and type of results it may take years. But during that time the work is improved, the whole process time can not be regarded is a 'delay' or a 'waste of time'. Furthermore, the average duration of a major research project is about 4 years, and the same is the case for most peer review time cycles. Also, during the publication process the awareness of scientific community (and peers!) evolves (e.g., average time between field–specific conferences etc.). We also have cases where the analysis can be performed almost in 'real time', as illustrated by an example[5] of a recent physics paper with citing articles published in the same year as the cited publication.

[5] Publication in *Physical Review Letters*, vol. 88, page 138701, year of publication 2002. The first citing articles are in the same year as the cited publication, we show the first four:
Marc Barthélemy *et al.,* Phys. Rev. E **66**, 056110 (2002);
Petter Holme, Phys. Rev. E **66**, 036119 (2002);
Holger Ebel *et al.,* Phys. Rev. E **66**, 035103 (2002;
Haijun Zhou, Phys. Rev. E **66**, 016125 (2002).

The above implies that ‘bibliometric awareness’ does not necessarily take more time than ‘peer awareness’. Moreover, the bibliometric system itself proves empirically the robustness of the method simply by showing that in many cases indicators, based on citation analysis, for universities, institutes, and larger research groups, are remarkably stable, as illustrated by the results presented in Table 1.1. We conclude that recent past performance is a reliable predictor for near future performance.

We also have to keep in mind that the importance of a publication does not necessarily appear immediately, even to peers, and that identification of quality may take considerable time (Garfield, 1980). An interesting phenomenon in this respect is the ‘Sleeping Beauty in Science’, a publication that goes unnoticed (‘sleeps’) for a long time and then, almost suddenly, attracts a lot of attention (‘is awakened by the prince’). Recently the first extensive measurements of ‘delayed recognition papers’ (Glänzel et al., 2003) and the occurrence of Sleeping Beauties in the science literature (van Raan, 2004) have been reported. In the latter work an ‘awakening’ probability function is derived from the measurements, and the ‘most extreme Sleeping Beauty up to now’ identified.

2.3.3 Comparability of the different research systems

It is often quite problematic to understand and ‘unravel’ the structure of a research organisation in terms of ‘real’ units such as departments or research groups. There are major differences in research systems between countries. For instance, the University of London is no longer a university in the usual sense. It is an ‘umbrella organisation’ covering several different virtually autonomous universities. In Paris and other French cities no such umbrella structure exists, there we deal with completely autonomous universities which were originally part of one ‘mother university’. As a consequence it is very cumbersome to distinguish between departments of these different universities within a city. The two ‘Free Universities’ of Brussels (Vrije Universiteit Brussel, VUB, and the Université Libre de Bruxelles, ULB) are a notorious example in this sense. Another well known problem is the ‘interwovenness’ of the French CNRS and French universities.

This problem is, in fact, a ‘fine structure’ problem: matching bibliometric data (‘external’) with the ‘real fine structure’ (‘internal’) of the principal organisation (e.g., a university). In order to do this, we need accurate ‘fine-structure’ data per organisation. Moreover, this internal structure is ‘dynamic’: new departments, schools, and certainly new research groups are created all the time.

I see at least two possibilities for tackling this problem. The first is the ‘narrowing down of fields’: *the smaller the bibliometric ‘refinement’ of* *fields* (e.g., from neuroscience as a whole to brain infarct research as a specific research theme within neuroscience), *the more we approach 'real' units* such as research groups within the internal structure of a principal organisation: 'convergence principle'. The bibliometric mapping methodology discussed in Section 3 (and a detailed discussion by Noyons (2004) in this handbook) is particularly suited to this approach.

A second approach concerns networks of co-operating scientists: the analysis of collaborating researchers provides the *internal structure* of that specific (sub-)field in terms of co-authors. Thus the real 'working floor' groups are identified (Vinkler, 1993; Melin and Persson, 1996; Glänzel, 2001). This identification is completely *independent* of the quality of information about principal organisation addresses. It is, as it were, based on a 'bibliometrically driven' *self-organisation* of science.

More generally, the understanding of research systems would benefit from the integration of bibliometric and other scientometric indicators into sociologically oriented studies (Gläser and Laudel, 2001).

2.3.4 Statistical issues: general ones and some related to journal impact

Standard statistical techniques relate to quantities that are distributed approximately 'normally'. Many characteristics of research performance, particularly those based on citation analysis, are not normally, but very skewly, distributed. Thus statistical averages can be misleading. For larger samples, such as the entire oeuvre of a research group over a period of years, the *central limit theorem* says that whatever the underlying distribution of a set of independent variables (provided that their variance is finite), the sum or average of a relatively large number of these variables will be a random variable with a distribution close to normal.

On the basis of these considerations I am confident that, for instance, our crown indicator ***CPP/FCSm*** does provide a useful measure. This can be proved empirically by the strong correlation of ***CPP/FCSm*** and the earlier discussed 'top 10%' indicator in which the distribution function is taken into account (Noyons et al., 2003).

A heavily debated theme in bibliometric studies is the 'predictive' character of journal impact, i.e., the relation between journal impact and the impact of a publication within that journal (see for instance Seglen, 1992, 1994; van Raan, 2001). In current research we focus in more detail on the relation between ***CPP*** and ***JCSm*** and other statistical characteristics of journal impact.

The indicators ***JCS*** and ***JCS/FCSm*** are novel journal indicators which characterise a journal in a more appropriate way than the commonly used journal impact factors. The unique aspect of these journal impact indicators is that the type of paper (e.g., letters, normal article, review) *as well as* the specific years in which the papers were published are taken into account. This is absolutely necessary, as the average impact of journals may have considerable annual fluctuations and large differences per article type, see Moed and van Leeuwen (1995, 1996).

2.3.5 Peer review judgment and bibliometric findings….signs of theory–invariance?

The results of peer review judgment and those of bibliometric assessment are not completely independent variables. Peers take 'bibliometric aspects' into account in their judgment, for instance (number of) publications in the better journals. Thorough studies of larger-scale evaluation procedures in which empirical material is available with data on both peer judgment as well as bibliometric indicators are rare. I refer to Rinia et al. (1998) for a comparison of bibliometric assessment based on various indicators with peer review judgment in condensed matter physics, and to Rinia et al. (2001) for a study of the influence of interdisciplinarity on peer review in comparison with bibliometric assessment.

I have already mentioned the empirical gold mine we created with our long standing bibliometric practice. In current work the relation between bibliometric assessment and peer judgment for several hundreds of physics and chemistry research groups is studied. This is a unique collection of data. This study shows a striking agreement between elements of research performance measurement and the results of peer review. But at the same time remarkable differences are found in which not necessarily peer judgment has to be considered as 'right' (van Raan en van Leeuwen, 2004).

Indeed, peers may be right or wrong in their judgement. Also they undoubtedly use bibliometric elements in their judgement; for instance, they generally attach great value to publications in the top journals. Therefore, bibliometric findings and outcomes of peer review are not independent variables in the 'quality judgment space'. But this entanglement is unavoidable because (1) there is no higher authority to judge the quality of scientific work than a peer group of colleagues, and (2) attracting attention, provoking reactions by written communication, is very fundamental in most fields of science. *Any reasonable theory* has to 'accept this reality'. So if bibliometric analysisis is advanced in such a way that it becomes an indispensable instrument for measuring progress of science, and we think this stage is reached now, then we are approaching Holton's ideal of 'theory–invariant' indicators.

3. PRINCIPLES OF CONCEPT–SIMILARITY BASED MAPPING

Each year about a million scientific articles are published. How should one keep track of all these developments? Are there specific patterns 'hidden' in this mass of published knowledge at a 'meta level', and if so, how can these patterns be interpreted (Van Raan and Noyons, 2002)?

A first and crucial step is the definition of a research field. There are several approaches: on the basis of selected concepts (keywords) and/or classification codes in a specific database, selected sets of journals, a database of field–specific publications, or any combination of these approaches. Along these lines titles and abstracts of all relevant publications can be collected for a series of successive years, thus operating on many tens of thousands of publications per field. Next, with a specific computer-linguistic algorithm, titles and abstracts of all these publications can be parsed. This automated grammatical procedure yields all nouns and noun phrases (standardised) which are present in the entire set of collected publications (Noyons, 1999).

An additional algorithm creates a frequency list of these many thousands of parsed nouns and noun phrases while filtering out general, trivial words. The most frequent nouns/noun phrases can be considered as the most characteristic concepts of the field (this can be 100 to 1,000 concepts, say, N concepts). The next step is to encode each of the publications with these concepts. In fact this code is a binary string (yes/no) indicating which of the N concepts is present in title or abstract. This encoding is as it were the 'genetic code' of a publication. As in genetic algorithms, the encoding of each publication can be compared with that of any other publication by calculating pairwise the 'genetic code similarity' (here: concept similarity) of all publications in a specific field. The more concepts two publications have in common, the more these publications are related on the basis of concept similarity, and thus they can be regarded as belonging to the same sub-field, research theme, or research specialty. To use a biological metaphor: the more specific DNA elements two living beings have in common, the more they are related. Above a certain similarity threshold they will belong to a particular species.

The above procedure allows clustering of information carriers — the publications — on the basis of similarity in information elements — the concepts ('co-publication' analysis). Alternatively, the more specific concepts are mentioned together in different publications the more these concepts are related. Thus information elements are clustered ('co-concept' analysis). Both approaches, the co-publication and the co-concept analysis, are related by the rules of matrix algebra. In practice the co-concept approach (Noyons and Van Raan, 1998) is most suited to science mapping, i.e., the 'organisation of science according to concepts'.

Intermezzo: For a supermarket 'client similarity' on the basis of shopping lists can be translated into a clustering either of the clients (information carriers, in which the information elements are the products on their shopping lists) or of the products. Both approaches are important: the first gives insight into groups of clients (young, old, male, female, different ethnic groups, etc.); and the second is important for the spatial division of the supermarket into product groups.

In outline the clustering procedure is as follows. First, for each field a matrix is constructed which composed of co-occurrences of the ***N*** concepts in the set of publications for a specific period of time. This 'raw co-occurrence' matrix is normalised in such a way that the similarity of concepts is no longer based on the pairwise co-occurrences but on the co-occurrence 'profiles' of the two concepts in relation to all other concepts. This similarity matrix is the input for a cluster analysis. Standard hierarchical cluster algorithm including statistical criteria can be used to find an optimal number of clusters. The identified clusters of concepts represent in most cases recognisable 'sub-fields' or research themes. Each sub-field represents a sub-set of publications on the basis of concept–similarity profiles. If any of the concepts is in a publication, this publication will be attached to the relevant sub-field. Thus publications may be attached to more than one sub-field. This overlap between sub-fields in terms of joint publications is used to calculate a further co-occurrence matrix, now based on sub-field publication similarity.

To construct a map of the field, the sub-fields (clusters) are positioned by multi-dimensional scaling. Thus sub-fields with a high similarity are positioned in each other's vicinity, and sub-fields with low similarity are distant from each other. The size of a sub-field (represented by the surface of a circle) indicates the share of publications in relation to the field as a whole. A two-dimensional structure is not sufficient to cover all relations embedded in the underlying matrix. Particularly strong relations between two individual sub-fields are indicated by a connecting line.

A next step (Noyons et al., 1999) is the integration of mapping *and* performance assessment. It enables us to position actors (such as universities, institutes, R&D divisions of companies, research groups) on the worldwide map of their field, and to measure their influence in relation to the impact-level of the different sub-fields and themes. Thus a strategic map is created: who is where in science, and how strongly?

A series of maps of successive time periods reveals trends and changes in structure, and even may allow 'prediction' of near-future developments by extrapolation. Such changes in maps over time (field structure, position of actors) may indicate the impact of R&D programmes, particularly in research themes around social and economic problems. In this way our mapping methodology is also applicable in the study of the socio-economic impact of R&D.

Bibliometric maps provide an instrument which can be used optimally in an electronic environment. Moreover, there is a large amount of detailed information 'behind the maps'. Hence it is of crucial importance that this underlying information, particularly about research performance, can be retrieved in an efficient way, to provide the user with a possibility of exploring the fields and of judging the usefulness of maps against the user's own expertise. Advanced internet–based user-interface facilities are necessary (Noyons, 1999; Noyons, 2004, in this Handbook) to enable this further exploration of the maps and of the data 'behind the maps'. Thus bibliometric maps and their internet–based user-facilities will enable users to compare the scientific performance of groups/institutes with other 'benchmark' institutes. Likewise, the maps can be used for the selection of benchmark institutes, for instance institutes chosen by the experts.

Co-citation analysis provides an alternative type of mapping, but it unavoidably depends on the availability of citation (reference) data and thus its applicability is less general than concept–similarity mapping. Co-citation maps are based on the number of times two particular articles are cited together in other articles. The development of this analytical technique is based on the pioneering work of Henry Small (Small, 1973; Small and Sweeney, 1985; Small et al., 1985). When aggregated to larger sets of publications, co-citation maps indicate clusters of related scientific work (i.e., based on the same publications, as far as reflected by the cited literature). These clusters can often be identified as 'research specialties' (McCain, 1990; Bayer et al., 1990; White and McCain, 1998; Small, 1999; Prime et al., 2002). Their character may, however, be of a different kind compared with co-word based clusters: because they are based on citation practices they may reflect cognitive as well as social networks and relations (Braam et al., 1991a,b). Moreover, citations only reflect a part of the intellectual structure, and they are subject to a certain, often field–specific, time lag. For recent work on co-citation analysis for mapping research themes of socio-economic importance I refer to Schwechheimer and Winterhager (2001).

As Derek de Solla Price formulated twenty five years ago: "scientific papers themselves form a system with a visible structure and, indeed, one that appears highly deterministic: the universe of scientific papers exhibits a clustering structure in a space of surprisingly small dimensionality: most of the behaviour can be accounted for in the usual two dimensions of a geographical map. The clusters correspond remarkably well to entities that we intuitively feel to be the basic sub-fields of which science is composed. Whatever their physical reality, maps of science are certainly useful as heuristic tools." (Price, 1978).

Mapping of science is a fascinating endeavour. For a detailed discussion of important new developments in bibliometric mapping I refer to the contribution of Noyons (2004) in this Handbook.

4. CONCLUDING REMARKS AND OUTLOOK

The quantitative study of science aims at the advancement of our knowledge on the development of science, also in relation to technological and socio-economic aspects. Bibliometric methods play an important role in this field of research. The field is both problem oriented as well as basic in nature. There are important interdisciplinary links with philosophy, history and sociology of science, with policy and management studies, with mathematics and physics, and particularly with information science.

I distinguish four inter-related research themes: (1) the development of methods and techniques for the design, construction, and application of quantitative indicators on important aspects of science; (2) the development of information systems about science; (3) the study of the interaction between science and technology; and (4) the study of cognitive and socio-organisational processes in the development of scientific fields.

The work in the *first* research theme concerns empirical studies on the assessment of research performance and directly related aspects such as publication and citation behaviour, notions of scientific quality, differences in communication practices in the different disciplines, comparison with qualitative judgments by peers. Standardisation of indicators including analysis of citing papers to assess aspects of 'knowledge users' mark the development of the 'second generation' bibliometric analysis (Van Leeuwen, 2004). At the same time it will be of crucial importance to monitor the influence of the various forms of electronic publishing on all bibliometric indicators, ranging from the mere number of publications to composed indicators such as the internationally normalised impact.

It is interesting to notice that only recently, owing to the gradually increasing number of applications of large-scale bibliometric analysis for research performance assessment, bibliometric characteristics of 'real' working floor entities such as research groups become known. So far, these characteristics have mainly concerned 'standard entities' such as authors, journals, universities, and countries. The study of the 'real working floor' enables the inclusion of further input data about personnel which goes beyond the data which are strictly necessary for conducting the bibliometric analysis described in Section 2.2. For instance, data about the sex and age of researchers enables one to investigate the role of women (Lewison, 2001; Prpić, 2002) or of the different age categories in the science system.

We have emphasised in this contribution the potential of advanced bibliometric indicators as 'theory–invariant' measures of scientific progress. Nevertheless, in the application of bibliometric indicators, no matter how advanced, it will remain of the utmost importance to know the limitations of the method and to guard against misuse, exaggerated expectations of non-expert users, and undesired manipulations by scientists themselves (Adam, 2002; Butler, 2003; Weingart, 2003; Glänzel and Debackere, 2003).

Given the crucial role of data as building blocks for indicators, it is not a surprise that a considerable part of the research in the field is devoted to the *second* theme: the development and maintenance of science information systems. These systems may contain data of many millions of scientific publications, but equally important are the many methodological and technical 'added values'. This part of quantitative studies of science is mainly system design and software development, in order to handle the enormous data system and to apply complex algorithms for the calculation of a wide range of indicators, including new journal impact measures. In addition, other than the 'classic' bibliometric data may be added to enrich the system with, for instance, input data of scientific institutions and business companies, patent data, and web-based data (Björneborn and Ingwersen, 2001; Bar-Ilan, 2001; Thelwall and Harries, 2003). Here we have an interdisciplinary bridge to information and computer science.

In the *third* research theme the focus is on the interaction between science and technology. I mention as an example the study of author–inventor relations (i.e., scientists who are active both in writing research publications as well as in creating technological breakthroughs), and the use of scientific knowledge in technological innovations (Schmoch, 1993) on the basis of citation relations between patents and publications (Albert et al., 1991; Narin, 1994; Narin et al., 1997; Glänzel and Meyer, 2003). Technology in its turn strongly influences scientific progress (Etzkowitz and Leydesdorff, 2000), particularly by the ever advancing development of instruments and facilities. Therefore the study of the interaction between science and technology has to take a broader perspective than only the transfer of knowledge from science to the technological domain. Most probably the development of instruments is *the* driving force of science. Hence the development of indicators describing the 'instrumental state-of-the-art' in scientific fields is very important.

The *fourth* theme is strongly related to bibliometric mapping techniques. The central issue here is to find optimal visual representations at different aggregation levels by exploring the idea of 'self-organising structures' in scientific and technological (on the basis of patents) development. It is a challenge to identify 'hidden patterns' in the enormous amount of data because all these publications (and patents) are connected by common references, concepts, classification codes. Co-citation and co-word techniques are examples of approaches to unravelling this gigantic network of inter-related pieces of scientific knowledge. These are important steps toward imaging cognitive processes. Systematic comparison of *cognitive* structures with *communication* structures based on citation analysis (Van Raan and Noyons, 2002) offers the possibility of discovering areas of science which are cognitively related but not connected in terms of reference practices (pioneering work by Swanson, 1986 and 1987).

Maps of science, with the locations of the major actors, are specific representations of scientific activities. They have practical values ('strategic overviews') as well as more cognitive (e.g., what type of scientific activities are primarily represented on the map). Co-word (concept similarity based) clusters can be used as '*journal set' independent* entities for defining (sub-) fields and research themes. An important advance in mapping is 'real time' user–driven application. This enables us to observe how differences in the definitions of fields (in terms of keywords, journals, etc.) lead to different maps, and, particularly, which defining elements really do matter. It also allows simulations and other manipulations that may teach us more about the meaning of science maps. This real-time mapping is absolutely necessary for making the next step: to know more about the relation between *cognitive* and *bibliometric* mapping.

Finally, an exciting development is the study of statistical and topological properties of bibliometric networks and their relation to other networks. Theoretical work is oriented towards the understanding of fractal properties of science as a 'bibliometric structure' in general, and of co-occurrence structures such as found in maps based on co-citation analysis in particular. Most probably these properties are related to (cumulative) growth phenomena (van Raan, 1990, 2000b). Soon the mapping and the network–based approaches will amalgamate. Bibliometric analysis then will reach its ultimate goal: to become, in the first place, an instrument for a scientist as a *grateful user*, instead of an instrument for a scientist as a *vulnerable target*.

To conclude this contribution, it is now not too vain to answer Holton's major question 'Can science be measured?' with a modest 'yes'.

REFERENCES

Adam, D. (2002). The counting house. *Nature,* 415, 726–729.

Albert, M.B., Avery, D., Narin, F., MacAllister, P. (1991). Direct validation of citation counts as indicators of industrially important patents. *Research Policy,* 20, 251–259.

Arunachalam, S., Srinivasan, R., Raman, V. (1994). International collaboration in science-participation by the Asian giants. *Scientometrics,* 30, 7–22.

Bar-Ilan, J. (2001). Data collection methods on the Web for informetric purposes — A review and analysis. *Scientometrics,* 50, 7–32.

Bayer, A.E., Smart, J.C., McLaughlin, G.W. (1990). Mapping intellectual structure of a scientific subfield through author cocitations. *Journal of the American Society for Information Science,* 41, 444–452.

Beaver, D. de B., Rosen, R. (1978). Studies in scientific collaboration, 1: Professional origins of scientific co-authorship. *Scientometrics,* 1, 65–84.

Björneborn, L., Ingwersen, P. (2001). Perspectives of webometrics. *Scientometrics,* 50, 65–82.

Borgman, C.L. (ed.) (1990). *Scholarly Communication and Bibliometrics*. Newbury Park: Sage.

Braam, R.R., Moed, H.F., van Raan, A.F.J. (1991a). Mapping of science by combined co-citation and word analysis, I: Structural Aspects. *Journal of the American Society for Information Science (JASIS),* 42, 233–251, and II: Dynamical Aspects. *Journal of the American Society for Information Science (JASIS),* 42, 252–266.

Braun, T., Glänzel, W., Schubert, A. (1988). World flash on basic research — The newest version of the facts and figures on publication output and relative citation impact of 100 countries 1981–1985. *Scientometrics,* 13, 181–188.

Braun, T., Glänzel, Grupp, H. (1995). The scientometric weight of 50 nations in 27 science areas, 1989–1993. 1: All fields combined, mathematics, engineering, chemistry and physics. *Scientometrics,* 33, 263–293; and 2: Life sciences. *Scientometrics,* 34, 207–237.

Brooks, T.A. (1986). Evidence of complex citer motivations. *Journal of the American Society for Information Science,* 37, 34–36.

Butler, L. (2003). Modifying publication practices in response to funding formulas. *Research Evaluation,* 17, 39–46.

Callon, M., Bauin, S., Courtial, J.P., Turner, W. (1983). From translation to problematic networks: an introduction to co-word analysis. *Social Science Information,* 22, 191–235.

Cole, S., Cole, J.R., Dietrich, L. (1978). Measuring the cognitive state of scientific disciplines. In: Elkana et al., op. cit.

de Candolle, A. (1873, 2nd. edition 1885). *Histoire des sciences et des savants depuis deux siècles.* Genève/Basel: H.Georg. Reprint in 1987 by Fayard.

Egghe L., Rousseau, R. (2000). The influence of publication delays on the observed aging distribution of scientific literature. *Journal of the American Society for Information Science,* 51, 158–165.

Elkana, Y., Lederberg, J., Merton, R.K., Thackray, A., Zuckerman, H. (Eds.) (1978). *Toward a metric of science: the advent of science indicators*. New York: John Wiley.

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.

Garfield, E. (1979). Is citation analysis a legitimate evaluation tool? *Scientometrics,* 1, 359–375.

Garfield, E. (1980). Premature discovery or delayed recognition — Why? *Current Contents,* 21, May 26, 5–10.

Gilbert, G.N. (1978). Measuring the growth of science– review of indicators of scientific growth. *Scientometrics,* 1, 9–34.

Glänzel, W. (1996). A bibliometric approach to social sciences, national research performances in 6 selected social science areas, 1990–1992. *Scientometrics,* 35, 291–307.

Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics,* 51, 69–115.

Glänzel, W., Meyer, M. (2003). Patents cited in the scientific literature: An exploratory study of 'reverse' citation relations. *Scientometrics,* 58, 415–428.

Glänzel, W., Schlemmer, B., Thijs, B. (2003). Better late than never? On the chance to become highly cited only beyond the standard bibliometric time horizon. *Scientometrics,* 58, 571–586

Glänzel, W., Debackere, K. (2003). *On the opportunities and limitations in using bibliometric indicators in a policy relevant context*. In: Bibliometric analysis in science and research. Applications, Benefits and Limitations. Second Conference of the Central Library, Forschungszentrum Jülich, (pp. 225– 236). (ISBN 3-89336-334-3).

Gläser, J., Laudel, G. (2001). Integrating scientometric indicators into sociological studies: methodical and methodological problems. *Scientometrics,* 52, 411–434.

Grupp, H., Schmoch, U., Hinze, S. (2001). International alignment and scientific regard as macro-indicators for international comparisons of publications. *Scientometrics,* 51, 359–380.

Haitun, S.D. (1982). Stationary scientometric distributions. 1: Different approximations. *Scientometrics,* 4, 89–104.

Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics,* 44, 193–215.

Holton, G. (1978). *Can science be measured*? In: Elkana et al., op. cit.

Horrobin, D.F. (1990). The philosophical basis of peer review and the suppression of innovation. *Journal of the American Medical Association (JAMA),* 263, 1438–1441.

Kamerlingh Onnes, H. (1882). *De betekenis van kwantitatief onderzoek in de natuurkunde* (The meaning of quantitative research in physics). Inaugural Address as Professor of Physics, Leiden University.

Koenig, M.E.D. (1983). Bibliometric indicators versus expert opinion in assessing research performance. *Journal of the American Society for Information Science,* 34, 136–145.

Kostoff, R.N. (1995). Federal research impact assessment — Axioms, approaches, applications. *Scientometrics,* 34, 163–206.

van Leeuwen, Th.N., Moed, H.F., Tijssen, R.J.W., Visser, M.S., van Raan, A.F.J. (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics,* 51, 335–346.

van Leeuwen, Th.N. (2004). *Second generation bibliometric analysis*. Ph.D. Thesis Leiden University.

Lewison, G. (2001). The quantity and quality of female researchers: a bibliometric study of Iceland. *Scientometrics,* 52, 29–43.

Lewison, G. (2002). Researchers' and users' perceptions of the relative standing of biomedical papers in different journals. *Scientometrics,* 53, 229–240.

Lotka, A.J. (1926). The frequency distribution of scientific productivity. *J. Washington Acad. Sci.*, 16, 317–323.

MacRoberts, M.H., MacRoberts, B.R. (1996). Problems of citation analysis. *Scientometrics,* 36, 435–444.

MacRoberts, M.H., MacRoberts, B.R. (1988). Author motivation for not giving citing influences — A methodological note. *Journal of the American Society for Information Science,* 39, 432–433.

Martin, B.R., Irvine, J. (1983). Assessing basic research: some partial indicators of scientific progress in radio astronomy. *Research Policy,* 12, 61–90.

May, R.M. (1997). The scientific wealth of nations. *Science,* 275, 793–796.

McCain, K.W. (1984). Longitudinal author cocitation mapping — The changing structure of macroeconomics. *Journal of the American Society for Information Science,* 35, 351–359.

McCain, K.W. (1990). Mapping authors in intellectual space — A technical overview. *Journal of the American Society for Information Science,* 41, 433–443.

Melin, G., Persson, O. (1996). Studying research collaboration using co-authorships. *Scientometrics,* 36, 363–377.

Moed, H.F., van Leeuwen, Th.N. (1995). Improving the accuracy of the Institute for Scientific Information's Journal Impact Factors. *J. of the American Society for Information Science (JASIS)*, 46, 461–467.

Moed, H.F., van Leeuwen, Th.N. (1996). Impact factors can mislead. *Nature,* 381, 186.

Moed, H.F., Luwel, M., Nederhof, A.J. (2002). Towards research performance measurement in the humanities. *Library Trends,* 50, 498–520.

Moravcsik, M.J. (1975). Phenomenology and models of growth of science. *Research Policy,* 4, 80–86.

Moravcsik, M.J., Murugesan, P. (1979). Citation patterns in scientific revolutions. *Scientometrics,* 1, 161–169.

Moxham, H., Anderson, J. (1992). Peer review. A view from the inside. *Science and Technology Policy*, February 1992, 7–15.

Narin, F. (1976). *Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity*. Washington D.C.: National Science Foundation.

Narin, F. (1978). Objectivity versus relevance in studies of scientific advance. *Scientometrics,* 1, 35–41.

Narin, F. (1994). Patent bibliometrics. *Scientometrics,* 30, 147–155.

Narin, F., Hamilton, K.S., Olivastro, D. (1997). The increasing linkage between US technology and public science. *Research Policy*, 26, 317–330.

National Science Board (1973). *Science Indicators 1972*. Washington DC: Government Printing Office.

Nederhof, A.J. (1988). *The validity and reliability of evaluation of scholarly performance*. In: A.F.J. van Raan (ed). (1988), Handbook of Quantitative Studies of Science and Technology (pp.193–228). Amsterdam: Elsevier/North-Holland, (ISBN 0-444-70537-6).

Noma, E. (1982). An improved method for analysing square scientometric transaction matrices. *Scientometrics,* 4, 297–316.

Noyons, E.C.M., van Raan, A.F.J. (1998). Monitoring scientific developments from a dynamic perspective: self-organized structuring to map neural network research. *J. of the American Society for Information Science and Technology (JASIST),* 49, 68–81.

Noyons, E.C.M., Luwel, M., Moed, H.F. (1999). Combining mapping and citation analysis for evaluative bibliometric purpose. A bibliometric study on recent development in micro–electronics. *Journal of the American Society for Information Science and Technology (JASIST),* 50, 115-131.

Noyons, E.C.M. (1999). *Bibliometric mapping as a science policy and research management tool.* Ph.D. Thesis Leiden University. Leiden: DSWO Press (ISBN 90-6695-152-4).

Noyons, E.C.M., Buter, R.K., van Raan, A.F.J., Schmoch, U., Heinze, T., Hinze, S., Rangnow, R. (2003). *Mapping excellence in science and technology across Europe* (Part

1: Life sciences, Part 2: Nanoscience and nanotechnology). Report to the European Commission by the Centre for Science and Technology Studies (CWTS), Leiden University, and the Fraunhofer Institute for Systems and Innovation Research (Fraunhofer-ISI), Karlsruhe.

Noyons, E.C.M. (2004). *Science Maps within in a Science Policy Context*. This Handbook.

OECD (1963). *The measurement of scientific and technological activities,* 'Frascati Manual', Paris: Organization for Economic Co-operation and Development (OECD).

Peritz, B.C. (1983). A classification of citation roles for the social sciences and related fields. *Scientometrics,* 5, 303–312.

Porter, A.L., Chubin. D.E. (1985). An indicator of cross-disciplinary research. *Scientometrics,* 8, 161–176.

De Solla Price, D.J. (1978). Toward a model for Science Indicators. In: Elkana et al., op. cit.

De Solla Price, D.J. (1981). The analysis of scientometric matrices for policy implications. *Scientometrics,* 3, 47–53.

Prime, C., Bassecoulard, E., Zitt, M. (2002). Co-citations and co-sitations: A cautionary view on an analogy. *Scientometrics,* 54, 291–308.

Prpić, K. (2002). Gender and productivity differentials in science. *Scientometrics,* 55, 27–58.

van Raan, A.F.J. (ed). (1988). *Handbook of Quantitative Studies of Science and Technology*. Amsterdam: Elsevier/North-Holland (ISBN 0-444-70537-6).

van Raan, A.F.J. (1990). Fractal dimension of co-citations. *Nature,* 347, 626.

van Raan, A.F.J. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics,* 36, 397–420.

van Raan, A.F.J. (1997). Scientometrics: State-of-the-Art. *Scientometrics*, 38, 205–218.

van Raan, A.F.J. (1998). In matters of quantitative studies of science the fault of theorists is offering too little and asking too much. *Scientometrics,* 43, 129–139.

van Raan, A.F.J. (2000a). The Pandora's box of citation analysis: measuring scientific excellence, the last evil? In: B. Cronin and H. Barsky Atkins (eds.). *The Web of Knowledge. A Festschrift in honor of Eugene Garfield.* Ch. 15, p. 301–319. Medford (New Jersey): ASIS Monograph Series, 2000 (ISBN 1-57387-099-4).

van Raan, A.F.J. (2000b). On growth, ageing, and fractal differentiation of science. Scientometrics 47, 347–362.

van Raan, A.F.J. (2001). Two–step competition process leads to quasi power–law income distributions. Application to scientific publication and citation distributions. *Physica A,* 298, 530–536.

van Raan, A.F.J., Noyons, E.C.M. (2002). *Discovery of patterns of scientific and technological development and knowledge transfer*. In W. Adamczak, A. Nase (Eds.), Gaining Insight from Research Information. Proceedings of the 6th International Conference on Current Research Information Systems, University of Kassel, August 29–31, 2002 (pp. 105– 112). Kassel: University Press, (ISBN 3-933146-844).

van Raan, A.F.J., van Leeuwen, Th.N. (2002). Assessment of the scientific basis of interdisciplinary, applied research. Application of bibliometric methods in nutrition and food research. *Research Policy,* 31, 611–632

van Raan, A.F.J. (2003). *Reference–based publication networks with episodic memories*. E–print ArXiv cond-mat/0311318.

van Raan, A.F.J. (2004). Sleeping Beauties in Science. *Scientometrics,* 59, 461–466.

van Raan, A.F.J., van Leeuwen, Th.N. (2004). *Statistical aspects of research group performance, journal impact, and peer judgement.* To be published.

Rinia, E.J., van Leeuwen, Th.N., van Vuren, H.G., van Raan, A.F.J. (1998). Comparative analysis of a set of bibliometric indicators and central peer review criteria. Evaluation of condensed matter physics in the Netherlands. *Research Policy,* 27, 95–107.

Rinia, E.J., van Leeuwen, Th.N., van Vuren, H.G., van Raan, A.F.J. (2001). Influence of interdisciplinarity on peer–review and bibliometric evaluations. *Research Policy,* 30, 357–361.

Rip, A., Courtial, J.P. (1984). Co-word maps of biotechnology — An example of cognitive scientometrics. *Scientometrics,* 6, 381–400.

Schmoch, U. (1993). Tracing the knowledge transfer from science to technology as reflected in patent indicators. *Scientometrics,* 26, 193–211.

Schwechheimer, H., Winterhager, M. (2001). Mapping interdisciplinary research fronts in neuroscience: a bibliometric view to retrograde amnesia. *Scientometrics,* 51, 311–318.

Schubert A., Glänzel, W. (1983). Statistical reliability of comparisons based on the citation impact of scientometric publications. *Scientometrics,* 5, 59–74.

Seglen, P.O. (1992). The skewness of science. *Journal of the American Society for Information Science,* 43, 628–638.

Seglen, P.O. (1994). Causal relationship between article citedness and journal impact. *Journal of the American Society for Information Science,* 45, 1–11.

Small, H. (1973). Co-citation in the Scientific Literature: A New Measure of the Relationship Between Publications. *Journal of the American Society for Information Science*, 24, 265–269.

Small, H., Greenlee, E. (1980). Citation context analysis of a co-citation cluster- recombinant DNA. *Scientometrics,* 2, 1980.

Small, H., Sweeney, E. (1985). Clustering the Science Citation Index using co-citations, I: A Comparison of Methods. *Scientometrics*, 7, 393–404.

Small, H., Sweeney, E., Greenlee, E. (1985). Clustering the Science Citation Index using co-citations, II: Mapping Science. *Scientometrics*, 8, 321–340.

Small, H. (1999). Visualizing science by citation mapping. *Journal of the American Society for Information Science,* 50, 799–813.

Swanson, D.R. (1986). Fish oil, Raynaud's syndrome, and undiscovered public knowledge. *Perspectives in Biology and Medicine,* 30, 7–18.

Swanson, D.R. (1987). Two medical literatures that are logically but not bibliographically connected. *Journal of the American Society for Information Science,* 38, 228–233.

Sullivan D., Koester, D., White, D.H., Kern, R. (1980). Understanding rapid theoretical change in particle physics- a month-by-month co-citation analysis. *Scientometrics,* 2, 309–319.

Thelwall, M., Smith, A. (2002). Interlinking between Asia-Pacific University Web sites. *Scientometrics,* 55, 363–376.

Thelwall, M., Harries, G. (2003). The connection between the research of a university and counts of links to its web pages: An investigation based upon a classification of the relationships of pages to the research of the host university. *Journal of the American Society for Information Science,* 54, 594–602.

Vinkler, P. (1993). Research contribution, authorship and team cooperativeness. Scientometrics 26, 213–230.

Vinkler, P. (1998). Comparative investigation of frequency and strength of motives toward referencing, the reference threshold model- comments on theories of citation? *Scientometrics,* 43, 107–127.

Vláchy, J. (1979). Mobility in science. Bibliography of scientific career migration, field mobility, international academic circulation and brain drain. *Scientometrics,* 1, 201–228.

Weingart, P. (2003). *Evaluation of research performance: the danger if numbers*. In: Bibliometric analysis in science and research. Applications, Benefits and Limitations. Second Conference of the Central Library, Forschungszentrum Jülich (pp. 7–19). (ISBN 3-89336-334-3).

White, H.D., Griffith, B.C. (1981). Author cocitation- a literature measure of intellectual structure. *Journal of the American Society for Information Science,* 32, 163–171.

White, H.D., McCain, K.W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972–1995. *Journal of the American Society for Information Science,* 49, 327–355.

Wouters, P.F. (1999), The Citation Culture, PhD thesis, University of Amsterdam.

Ziman, J. (1978). *From Parameters to Portents -and Back*. In: Elkana et al., op.cit.

Chapter 2

ECONOMETRIC APPROACHES TO THE ANALYSIS OF PRODUCTIVITY OF R&D SYSTEMS

Production Functions and Production Frontiers

Andrea Bonaccorsi[1] and Cinzia Daraio[2]

[1] *School of Engineering, University of Pisa, Italy*
[2] *Institute for Informatics and Telematics – CNR and Sant'Anna School of Advanced Studies, Pisa, Italy. E-mail: cinzia@sssup.it*

Abstract: In this chapter we review and discuss the potential and limitations of econometric methods for the evaluation of productivity of scientific and technological (S&T) systems. We examine and compare the main approaches that have been applied in the literature: the production function and the production frontier approach. Both approaches present advantages and disadvantages. In the first part of the chapter we carry out a selective review of the two fields. In the second part we focus on the last developments of the efficiency analysis literature, with particular attention to the nonparametric approach. An illustration of the potential of robust nonparametric techniques is offered using data from the Italian National Research Council (CNR). The chapter concludes by discussing the potential of these approaches for the analysis of S&T systems beyond the existing applications.

1. INTRODUCTION

In this chapter we review and discuss the potential and limitations of econometric methods for the evaluation of productivity of S&T systems.

Any notion of productivity relates a vector of inputs to a vector of outputs. Unfortunately, in S&T systems all three definitional elements of productivity (inputs, outputs and the functional relation between the two) are affected by severe conceptual and measurement problems.

S&T production is based on a multi-input, multi-output relation, in which, differently from standard production activity, both inputs and outputs

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 51-74.

are not only qualitatively heterogeneous but sometimes truly incommensurable, the relation between inputs and outputs is non-deterministic, and the output is lagged but with a lag structure which is not fixed.

The econometric approach to the analysis of R&D systems has taken two main directions. The former refers to the estimation of the structure of production of scientific and technological output by individual units (e.g., universities, research institutes, firms), the latter to the estimation of the impact of including science and technology as inputs in a more general production relation at the macroeconomic level. We will focus on the former types of problems, although we will take the latter into account in terms of the econometric problems which have been discussed and (sometimes) solved[1].

We examine and compare the main approaches that have been applied in the literature in order to deal with these problems: the production function approach and the production frontier approach (efficiency analysis).

In the production function approach the measurement of scientific productivity is carried out by specifying a functional relation which *intersects* observed data, looking for average relations, and estimating coefficients that relate inputs to outputs.

In the production frontier approach, the interest lies in estimating a frontier that *envelops* the datapoints and in measuring the distance between each observed unit and the estimated 'efficient' frontier.

With respect to the estimation of coefficients in the production function, the task of approximating the mean function can be done essentially in three ways. The *parametric* approach assumes that the mean curve has some pre-specified functional form, e.g., a line with unknown slope and intercept. As an alternative one could try to estimate the mean function *nonparametrically*, i.e., without reference to a specific functional form. Finally, one could choose an intermediate solution. In fact, using a *semiparametric* approach, a part of the model is parameterised and another part is not.

With respect to production frontiers, on the contrary, the estimation of efficiency indexes is made by comparing each unit with the best performers in the reference group. The best performers are defined as those units which obtain the maximum level of output given their level of inputs (the input oriented approach) or minimise the inputs utilised given the level of outputs obtained (the output oriented approach). By definition, an efficiency index

[1] For a survey of econometric studies that investigate the relationships between R&D and productivity, see Mairesse and Sassenou (1991). See also Hall and Mairesse (1995).

gives a score relative to another unit, without any reference to absolute efficiency.

Production frontiers can be estimated following parametric, nonparametric or semiparametric estimation methods. The former specify functional form for the frontiers that envelope observed datapoints, whilst nonparametric methods leave the determination of the shape of the envelope to the data itself. Again, the semiparametric estimation method combines the two.

2. A SELECTIVE REVIEW OF THE LITERATURE

The measurement of productivity in S&T systems can follow different strategies. In the following we give a description of the main approaches, starting with a brief outline of ratio measures and index numbers and describing more deeply the measures based on production functions and production frontiers.

A very simple approach is based on a crude comparison of *simple measures of productivity* expressed as output/input ratios. This approach takes one type of input and relates it to one type of output, ignoring all relations of complementarity and substitution between inputs, and all effects of joint production in outputs. They serve mainly as a first order approximation.

Ratios of output to input are clearly *partial* productivity measures. This terminology distinguishes them from *total* factor productivity measures because the latter try to obtain a value of the output to input ratio which takes into account *all* outputs and inputs. Moving from partial to total factor productivity measures by combining all inputs and all outputs to obtain a single ratio helps to avoid imputing gains to one factor (or one output) that should be attributed to some other input (or output). However, total factor productivity measures present aggregation problems such as choosing the weights to be used in order to obtain a 'single output to single input' ratio.

An index number is defined as a real number that measures changes in a set of variables. In particular, index numbers are applied to measure price and quantity changes over time, as well as to measure differences in the levels across firms, industries, regions, or countries. Panel data allow the measurement of productivity change as well as the estimation of technical progress or regress. Productivity change occurs when an index of outputs changes at a different rate from that at which an index of inputs does. Productivity change can be calculated using index number techniques such as Fischer or Tornqvist productivity indices. Both these indices require quantity and price information, as well as assumptions about the structure of the technology and the behaviour of producers.

Productivity change can also be calculated using a production frontier approach to construct a Malmquist productivity index. This approach does not require price information or technological and behavioural assumptions, and allows the identification of the sources of measured productivity change (i.e., technological progress/regress, and efficiency changes). It requires the estimation of a representation of production technology that can be made using both a parametric and a nonparametric frontier approach. A survey of the theoretical and empirical work on Malmquist productivity indices can be found in Färe, Grosskopf, and Russell (1998), while some applications to the efficiency and productivity of colleges and university licensing can be found in Førsund and Kalhagen (1999), Thursby (2000), Thursby and Kemp (2002), Thursby and Thursby (2002).

2.1 Production Functions

Theoretical mainstream production analysis focus on production activity as an optimisation process. On the other hand, empirical production analysis has focused on a central tendency, or 'average' or 'most likely' relationship constructed by intersecting data with a function.

Production functions are based on equations which relate quantities of inputs to quantities of outputs. More precisely, the production function is a mathematical function (a relation) which associates (relates) the vector of input X with the maximum level of output Y [2].

From the empirical point of view estimating production functions means estimating the coefficients of regression equations which describe the average tendency of the relationship between inputs and outputs. In production functions the notion of efficiency refers to the average behaviour, not the individual behaviour of each unit.

The production function framework applies to production process which are well specified, i.e., to well structured production processes.

[2] By means of its parameters, it is possible to analyse: the level of productivity, which is usually given by a coefficient which multiplies the function (this is the case of *neutral* technical progress); the marginal productivity of each factor (making the assumptions that the factors can be measured without ambiguity, the other inputs can be kept constant, the availability of an infinite number of techniques such that the passage from one combination of factors to another could happen also for infinitesimal variations); the marginal rate of substitutions amongst factors; the factors' intensity, given by the ratio of the amount of two inputs, given the marginal rate of substitutions; the optimal choice of the combination of inputs, trough the equality of the factors' marginal rate of substitutions and their prices ratio; simple measure of productivity by doing the ratio of the observed level of output over the production function optimal level; measures of technical change; returns to scale; inputs' elasticity of substitution.

In the field of S&T production functions have been used in both the estimation of production of scientific and technological output and in the estimation of the impact of S&T on economic growth.

Within the former line of research, Adams and Griliches (1998) used a Cobb Douglas specification to study the relation between funding and published output of American universities and to estimate the presence and magnitude of economies of scale at the level of university and Arora, David and Gambardella (1998) estimated the production function for scientific publications in the field of biotechnology. Several other functional forms have been introduced in the literature to describe the relation between inputs and outputs (useful reviews on the production function forms are Nadiri, 1970, and Heathfield and Wibe, 1987; a review of empirical findings about productivity is in Bartlesman and Doms, 2000).

Within the latter domain it is useful to recall the remark of Mairesse and Sassenou (1991), who pointed out that "most econometric studies that attempt to assess the contribution of R&D to economic growth rely on the Cobb Douglas production function as their basic analytical framework".

The adoption of a production function modeling strategy is based on a number of assumptions whose limits have been highlighted in the literature on the economics of education, but also apply to the domain of the economics of research.

First, it is normally assumed that the production function is homothetic, that is, "the marginal rate of substitution among inputs (…) depends only on the proportions of the inputs and not on the scale of production" (Figlio, 1999, p. 242). This means that the relative impact of the addition of one unit of any given input will be the same irrespective of the size of the output (Gyimah-Brempong and Gyapong, 1992; Hanushek, Rivkin, and Taylor, 1996). In the Cobb Douglas formulation *elasticity of substitution* (measuring the percentage change in factors' proportion owed to a change in marginal rate of substitution) is considered constant. Second, production functions require additivity of inputs, excluding interaction effects.

These assumptions may be considered restrictions within a more general specification, such as the translog or trascendental logarithmic (Griliches and Ringstad, 1971; see also Nadiri, 1970; Heathfield and Wibe, 1987). In particular, within this specification additivity requires that interaction terms are set to zero. Studies which adopt a more flexible specification generally conclude that the assumption of homotheticity is rejected (Nelson and Hevert, 1992; de Groot, McMahon and Volkwein, 1991).

In parallel, a consistent body of literature has worked with a multi-product cost specification based on the analysis of economies of scale and scope proposed by Baumol, Panzar, and Willig (1982). Using a flexible fixed cost quadratic function it is possible to take into account differences in fixed costs associated with different outputs, abandoning the linear homogeneity property of costs with respect to the prices of factors (Cohn, Rhine and Santos, 1989; de Groot, McMahon, Volkwein, 1991; Dunbar and Lewis, 1995; King, 1997). With this specification it is possible to estimate economies of scope with respect to all possible combinations of outputs and to the overall effect. Since research activities are intrinsically multi-output, the estimation of economies of scope is a critical issue, particularly with respect to the teaching research complementarity. As shown by Cohn et al. (1989) the use of multi-output cost functions may lead to qualitatively different results than with single output models.

Although these specifications are much more flexible than the standard Cobb Douglas, they still rely on a pre-specified functional form.

2.2 Production Frontiers

On the contrary, the approach of production frontiers (see, e.g., Färe, Grosskopf, and Lovell, 1994) is based on the envelopment of production data. From the empirical point of view it offers techniques for estimating the 'efficient' production frontier and for measuring and interpreting the relative efficiency of each individual unit with respect to this estimated frontier.

The purpose of efficiency analysis based on frontiers is to make a relative benchmark or comparison among decision making units (DMUs). Each DMU is compared to the best performer included in the analysis. The comparison is therefore made on the basis of the real or observed performance of units, and not the theoretical maximum as derived from a production function.

Nonparametric frontiers do not require the user to prescribe weights to be attached to each input and output, as in the usual index number approaches, and do not require prescribing the functional forms which are needed in regression approaches.

Efficiency measures are obtained by comparing each institute to the most efficient ones in its own comparison set. The most efficient institutes are those which minimise the use of inputs given a level of observable outputs (input oriented), or maximise outputs given a level of observable inputs (output oriented).

The structure of production frontiers can be different from the structure of production functions constructed from the same data. Best practice is not just better than average practice, it may also be structurally different, and it is important to know whether the structure of efficient production differs from the structure of average production. Best practice may be better in the sense that it exploits available substitution possibilities or scale opportunities that average practice does not. Public policy based on the structure of best practice frontiers may be very different from policy based on the structure of average practice functions.

This approach is more appropriate for production processes in which the variance of output may be extremely high, for example because of the skewness of the underlying distribution.

Efficiency analysis has been developed from the first empirical work of Farrell (1957) which defines a simple measure of firm efficiency which could account for multiple inputs and multiple outputs: “when one talks about the efficiency of a firm one usually means its success in producing as large as possible an output from a given set of inputs” (Farrell, 1957, p. 254). Farrell proposed that the efficiency of a firm consists of two components: *technical efficiency*, which reflects its ability to obtain maximal output from a given set of inputs, and *price* (or *allocative*) *efficiency*, which reflects the ability of a firm to use the inputs in optimal proportions, given their respective prices and the production technology. Starting from Farrell’s pioneering work mainly two approaches developed for the estimation of the ‘efficient frontier’:

a) A nonparametric approach based on the estimation of a piecewise linear convex frontier, constructed such that no observed point lies to the left or below it;
b) A parametric approach based on a function fitted through the data, such that no observed point lies to the left or below it.

Following point a), Charnes, Cooper, and Rhodes (1978) proposed the Data Envelopment Analysis (DEA) approach. DEA involves the use of linear programming methods to construct a non parametric piecewise surface (or frontier) over the data. It is based on the free disposability and convexity assumptions for the production set (the set of the attainable points). Free disposability means that the destruction of goods is not expensive. Convexity implies that the efficient frontier includes all linear combinations of dominant units.

A more general nonparametric approach is the Free Disposal Hull (FDH), introduced by Deprins, Simar, and Tulkens (1984). FDH assumes only the free disposability of the production set.

Efficiency measures are then calculated relative to this surface[3]. Elasticities, measuring the degree of substitutability between pairs of factors,

[3] Charnes, Cooper, and Rhodes (1978) proposed a model that had an input orientation and assumed constant returns to scale (CRS). In their original study they described DEA as a “mathematical programming model applied to observational data that provides a new way of obtaining empirical estimates of extreme relations such as the (*average*, n.o.w.) production

can be computed through the parametrization of the nonparametric frontiers. They do not describe average values but the shape of the frontier.

Returns to scale are estimated pointwise and globally. This allows one to track returns to scale in different regions of the size distribution. The analysis of efficiency indexes gives information on those inputs which are wasted (i.e., do not contribute to output) through the *analysis of slacks*.

From this original formulation an impressive literature developed, with a number of extensions and refinements. At present DEA encompasses a variety of models for evaluating performance[4]. A large literature has applied Data Envelopment Analysis to problems of productivity in a large number of manufacturing and service settings.

Several studies have used approaches of DEA type in assessing the efficiency of academic research, e.g., Johnes and Johnes (1993, 1995), Rizzi (1999), Korhonen, Tainio, and Wallenius (2001), Abbott and Doucouliagos (2003). Studies applying DEA to education include Bessent and Bessent (1980); Bessent, Bessent, Kennington and Reagan (1982); Charnes et al. (1978); Färe, Grosskopf, and Weber (1989); Thanassoulis and Dunstan (1994); Sarrico, Hogan, Dyson and Athanassopoulos (1997); Grosskopf, Hayes et al. (1999); Grosskopf and Moutray (2001); and Grosskopf et al. (2001).

Rousseau and Rousseau (1997, 1998) applied DEA to construct scientometric indicators and assess research productivity across countries. Bonaccorsi and Daraio (2003a) used DEA together with FDH and measures of order *m* to compare two large research institutions (CNR and INSERM) in different countries in the biomedical field.

The parametric approach was introduced by Aigner and Chu (1968) who developed the deterministic frontier model approach based on the estimation of a parametric frontier production function of Cobb Douglas form. This

functions and/or efficient production possibility surfaces that are a cornerstone of modern economics".

[4] Banker, Charnes, and Cooper (1984) proposed an extension of the CRS DEA model to account for variable returns to scale (VRS) situations. The Banker, Charnes and Cooper (1984) model distinguishes between technical and scale inefficiencies by estimating pure Technical Efficiency (TE) and the Scale Efficiency (SE). The TE is a measure of the radial distance of a unit to the estimated efficient frontier. If TE is equal to 1 then the decision unit is located on the efficient frontier. If TE is less than 1 (input oriented), its value represents the proportionate reduction of inputs (given the value of outputs) the unit should put in place, in order to be fully efficient. The SE can be roughly interpreted as the ratio of the average product of a unit to the average product of a unit operating at a point of technically and optimal scale. If it is 1 the DMU is scale efficient, if it is less than 1 the unit is scale inefficient.

approach is called deterministic because in the frontier model the observed output is bounded above by the non-stochastic deterministic quantity.

One of the main criticisms of the deterministic frontier model is that no account is taken of the possible influence of measurement errors and other noise upon the frontier. All deviations from the frontier are assumed to be the result of technical inefficiency. Aigner, Lovell, and Schmidt (1977) amongst others, proposed the stochastic frontier production function, in which an additional random error was added to the non-negative random variable which represents inefficiency. For a survey of recent contributions on the parametric frontier analysis, see Kumbhakar and Lovell (2000).

A multi-output specification within a parametric frontier approach was developed by Grosskopf, Hayes, Taylor, and Weber (1997) using the indirect output distance function initially proposed by Färe, Grosskopf, and Lovell (1988). Cooper and Cohn (1997) applied a parametric function and frontier approach to evaluate the productivity of the educational system of South Carolina.

More recently a semiparametric generalization of the parametric approach has been introduced in the literature. In this approach a part of the model is parametric and another part is nonparametric (for more details, see Park and Simar, 1994; Park, Sickes, and Simar, 1998; 2003).

Nonparametric production frontier techniques have several advantages for the analysis of S&T systems. Let us discuss them in detail.

2.2.1 Absence of specification

This property is particularly interesting for the analysis of S&T systems. Let us focus mainly on scientific production in the public sector research system. Scientific production is not only a multi-input multi-output process, but the relation between inputs and outputs is non-deterministic, uncertain, lagged, non-linear, and subject to important but subtle external effects.

We know from the economics of science (Stephan, 1996; Stephan and Levin, 1996) that a few stylized facts about individual productivity do exist. First, the distribution of individual productivity of scientists is extremely skewed, with a small percentage of very productive scientists accounting for a disproportionate share of publications. Second, productivity declines over a scientist's life cycle. These very basic features of scientific production make a representation in which the marginal rate of substitution between units of inputs is constant or independent on size, and in which interaction effects are zero, highly unrealistic.

How these individual level factors combine on an organizational and institutional level is, in fact, a very open question. Do people with the same individual productivity attract each other, or perhaps are hired according to a consistent quality strategy, so that in the end the same skewed distribution will also be observed across organizations and institutions? Or, quite to the contrary, do people with different individual productivities mix within research departments and institutes? What is the effect of the organizational setting on individual productivity?

External factors may create complementarities which have a non-linear effect. Studies of individual productivity of scientists (Fox, 1983; Holbrook, 1992; Johnston, 1993; Ramsden, 1994; Narin and Breitzman, 1995) often point to the extremely powerful effect of the external environment of scientists, in terms of complementary resources, time constraints, and social incentives at the level of department or institute.

Whilst these external effects are clearly important, it is difficult to capture them within a production function approach, above all a parametric one.

Under these conditions the lack of a specification is a clear advantage.

2.2.2 Aggregation of output indicators

Research activities are intrinsically multi-output activities.

First of all, for a large part of the research system the allocation of the time of researchers takes place between research and teaching. Since the share of time is not fixed across disciplines and countries, it is sensible to take both outputs into consideration, when possible.

Second, within the narrow area of research, whilst the single most important output is clearly scientific publications, it is difficult to claim that other outputs such as patents, software, advisory work for the government, consulting, or technical assistance do not have any relevance with respect to research.

Finally, scientific publications cover a large range of specific outputs, such as papers in refereed journals, papers in technical or professional journals, notes, reviews, books, and edited books. Even though, as in standard bibliometrics, one eliminates unpublished materials such as technical notes, working papers, and conference papers, there is still much heterogeneity. How much worth is a book with respect to a paper in a refereed journal? Do more papers in the technical press compensate for fewer papers in academic journals?

In order to take into consideration the multi-output nature of research it is necessary to aggregate each type of output. This may be done in two ways: assigning a weight to each type of output which is valid across all units of observation or using a multi-output specification.

The first solution has no alternative if one takes a production function approach based on a Cobb Douglas. The regression equation will have to be run on an independent variable that aggregates several types of outputs within a single measure. Owing to the lack of prices for most inputs and outputs of higher education and research, however, any weighting scheme which reflects their relative importance is fundamentally arbitrary[5]. More flexible forms such as translog allow the estimation of multi-input multi-output relations, but still under restrictive assumptions on the relations between inputs and outputs.

Nonparametric techniques radically solve the problem by allowing each unit to select the vector of weights which maximizes its own efficiency score. This is an interesting property for the analysis of S&T systems, whose evaluation is inevitably open to debate owing to its intrinsic heterogeneity and the impossibility of value–free statements about the hierarchy of outputs.

2.2.3 Pointwise estimation of efficiency

As has been illustrated before, nonparametric techniques allow the estimation of returns to scale and scope on each point of the interval. This is another interesting property for addressing a difficult issue in the economics of research, which has also a well developed counterpart in the economics of education.

In fact, there is lack of consensus on the existence of economies of scale in the production of research and university teaching. Amongst many others, Brinkman (1981), Brinkman and Leslie (1986), Cohn et al. (1989), de Groot, McMahon and Volkwein (1991), Nelson and Hevert (1992), and Lloyd, Morgan and Williams (1993) report the existence of economies of scale. Verry and Layard (1975), Verry and Davies (1976), and Adams and Griliches (1998), on the contrary, found constant returns to scale.

This problem has clear implications in terms of governmental policies (Bonaccorsi and Daraio, 2003c). For example, Abbott and Doucouliagos (2003) report that the Australian government, in the attempt to improve the efficiency of the university system by exploiting economies of scale and scope, consolidated a large number of higher education institutions into a small number of large multi-campus universities.

It is difficult to draw general implications from the existing evidence, mainly because data and methodologies are not strictly comparable.

Estimating economies of scale over the entire range of observations, as is standard in the production function, will result in *averaging* a number of very different *local* size effects. The policy implication of finding, for example,

[5] Some developments of DEA includes preference structure models (Zhu, 1996) where the target for inefficient DMUs is given by a preference structure (represented through some weights) expressed by the decision maker; and the value efficiency analysis (Halme et al., 2000) aims at incorporating the decision maker's value judgements and preferences into the analysis, using a two stage procedure.

economies of scale will be consolidating universities or merging research units. But if size effects are local the policy may even worsen the situation. Suppose there are several regions of returns to scale, initially increasing then constant or decreasing. Merging units means that smaller institutes, which initially benefited from economies of scale, will become larger and will enter into a region where these effects are eliminated.

On the contrary, in the nonparametric frontier approach it is possible to estimate separately the efficient frontier returns to scale, the global effect of scale, and the individual position with respect to returns to scale. As we shall see in the application at the end of this Chapter, it is possible that returns to scale are variable over a limited interval, whilst they are constant over other intervals of the observed size distribution.

The only way to give accurate policy implications will be to examine scale effects across the whole range of observations, paying attention to local effects. Techniques that estimate average returns to scale fail to identify all these effects.

2.3 Production Functions versus Production Frontiers in the Analysis of S&T Systems

In using production functions there are several interconnected methodological problems to be examined.

First of all, the problem of *identification* is crucial. Generally speaking, most empirical studies limit their task to describing the methodology of estimation and then interpret the obtained results. Before analysing the estimation and results, however, the fundamental issue of whether the parameters of interest in the model are even estimable must be resolved. (For an introduction to the problem of identification, see, e.g., Greene, 2000, pp. 663 ff. For an historical and detailed discussion see Griliches and Mairesse, 1998).

Second, *misspecification* concerns the problems and errors related to the assumptions made by the model. Empirically, misspecification errors are mainly related to the specification of explanatory variables, in particular, knowledge of which ones of the variables to include and about the mathematical form of their inclusions. A related topic is the exclusion of relevant variables and the inclusion of irrelevant variables. Policy making based on empirical evidence is strictly related to the assumptions of the econometric methodology applied. Several studies have largely discussed, for instance, the effect of misspecification in the evaluation of the performance of universities or schools (see Hanushek, 1986; Nelson and Hevert, 1992; Figlio, 1999; Pritcett and Filmer, 1999; Baker, 2001; Daneshvary and Clauretie, 2001; for a survey see Dewey, Husted and Kenny, 2000).

Third, the *simultaneity* in the relationship between variables could greatly affect the estimation of parameters creating a source of bias. This problem could be controlled for using a General Method of Moments (GMM) approach. GMM (for a general presentation, see Hansen, 1982) is a method for parameter estimation that can be viewed as a general case of OLS, instrumental variable estimation, two stage least squares, and so on. For an application of GMM to estimating the productivity of R&D see Hall and Mairesse (1996).

Finally, *multicollinearity* is the problem related to the existence of a linear dependence amongst the response or independent variables. The multicollinearity affects the problem of unidentifiability of the regression parameters.

A discussion of the hypothesis of the model and a *diagnostic* analysis on, e.g., the model residuals, are generally omitted in the studies we reviewed. As an example, autocorrelated residuals could be related to omitted variables, incorrect specification of the model, inter-temporal aggregation of the data, or incorrect specification of the error term.

Coefficients may be interpreted as elasticities of the output with respect to individual inputs. On the other hand, production functions do not allow the analysis of slacks of inputs.

It must be underlined, however, that even the adoption of all (sophisticated) techniques for improving the quality of the estimation of coefficients, or the adoption of a nonparametric regression approach, does not solve the fundamental problem of estimating the *expected* or *average* value.

This is appropriate for production process in which the variance of output is bounded around the average value. In S&T systems there is no *a priori* rationale that this is the case.

On the other hand, nonparametric frontier techniques also suffer from a number of limitations, although recently developments solve most of the problems.

A first limitation of the nonparametric approach in production frontier analysis is its *deterministic nature*. In this framework it is assumed that all deviations from the efficient frontier are owed to inefficiencies. The problem of handling noise in this context is owed to the model not being identified unless some restrictions are assumed. See, e.g., Aigner, Lovell, and Schmidt (1977) for approaches that assume a parametric function for the frontier; or Kneip and Simar (1996) for the case of panel data. More general results for handling noise in nonparametric frontier models can be found in Hall and Simar (2002) and in Simar (2003).

A second limitation of nonparametric techniques is the *more difficult economic interpretation* of the production process in terms of, e.g., shape of the production function, elasticities, etc. To overcome this drawback an alternative is represented by the analysis of slacks, that is, the excess resources wasted in the production activity (see, e.g., Färe, Grosskopf, and Lovell, 1994), whilst Florens and Simar (2002) propose the full theory for parametric approximations of nonparametric frontier.

The problem of *extremes* or *outliers* can be treated applying the recently introduced robust order *m* frontiers (Cazals, Florens and Simar, 2002). The order *m* frontiers represent a more realistic benchmark. Instead of comparing the performance of each unit with the best performers, the benchmark is done against the expected value of an appropriate sample of *m* units, drawn randomly from the population. The method offers flexibility in choosing the level of robustness of the estimate, by varying the parameter *m*.

The robust nonparametric frontiers of order-m do not suffer also from the so called 'curse of dimensionality'. Shared by many nonparametric methods the curse of dimensionality means that to avoid large variances and wide confidence interval estimates a large quantity of data is needed.

Zhang and Bartels (1998) show formally how DEA efficiency scores are affected by sample size. They demonstrate that comparing measures of structural inefficiency between samples of different sizes leads to biased results. This *sample size bias* problem can be easily overcome using the robust nonparametric approach based on order *m* frontiers.

Another limitation of the nonparametric approach is the *difficulty in making statistical inference*, owing to its complex nature: nonparametric estimation in a space at $p+q$ dimensions (where p is the number of the inputs and q is the number of the outputs), based on very few assumptions. Thanks to the last developments of the literature, statistical inference in nonparametric frontier models is available based on asymptotic results or on bootstrap application (for a review see Simar and Wilson, 2000). Asymptotic results are potentially useful for estimating asymptotic bias and variance, as well as asymptotic confidence intervals, but they remain asymptotic results which may be misleading in conjunction with small samples. Moreover, additional noise is introduced when estimates of the unknown parameters of the limiting distributions are used in constructing estimates of confidence intervals. Hence an attractive alternative to asymptotic results is represented by the bootstrap[6].

Useful bootstrap applications in a frontier analysis framework include the correction for the bias and the construction of confidence intervals for efficiency scores; applications to Malmquist indices and their various decompositions (see Simar and Wilson, 1999); tests procedure to assess

[6] The essence of the bootstrap idea is to approximate the sampling distributions of interest by simulating (or mimicking) the Data Generating Process. For an introduction to the bootstrap see Efron and Tibshirani (1993).

returns to scale (Simar and Wilson, 2002); statistical tests to compare the means of several groups of producers (see Simar and Zelenyuk, 2003).

In addition, there may be uncertainty about the structure of the underlying statistical model in terms of whether certain variables are relevant or whether subsets of variables may be aggregated. Tests of hypotheses about the model structure have been introduced (see Simar and Wilson 2001 for more details).

Finally, the traditional two stage approach used in nonparametric frontier models to explain efficiency scores relies on a second regression–based step which, as pointed out by Simar and Wilson (2003a), suffers from several problems. Daraio and Simar (2003) propose a probabilistic approach for evaluating the influence of external environmental variables that overcomes most drawbacks of previous approaches.

A summary of differences between production functions and frontiers is offered in Table 2.1.

Table 2.1. (Parametric) Production functions vs. (nonparametric) production frontiers

	Production functions	*Production frontiers*
Nature of production process	Well specified	Not specified
Functional specification	Yes	No
Estimation problems	Yes (identification)	Yes (curse of dimensionality)
Object of the estimation	Conditional expected value	Envelope
Economic interpretation	Parameters (elasticity)	No parameters
Returns to scale	Average effects	Pointwise and globally

3. A ROBUST AND PROBABILISTIC APPROACH TO EVALUATE AND EXPLAIN S&T PERFORMANCE

3.1 Some Basic Concepts

In this section we briefly outline the main ideas of a recently introduced probabilistic and robust nonparametric methodology for evaluating and explaining the productivity/efficiency of DMUs.

In the light of our previous discussion about the advancements of the nonparametric approach in frontier analysis we believe it is a promising approach to be used in the evaluation and explanation of the performance of S&T systems.

It is based on the probabilistic approach proposed by Daraio and Simar (2003) to explain the efficiency of production units. It relies on the concept of order *m* frontiers introduced by Cazals, Florens and Simar (2002), known as robust estimator of the efficient frontier, and applied to the evaluation of scientific productivity by Bonaccorsi and Daraio (2003a).

This methodology measures the productivity levels using a *nonparametric* production frontier approach that does not require the specification of any functional form for the production frontier. In particular, it has been implemented in a FDH framework that, with respect to a DEA context, assumes only the free disposability of the production set (and not its convexity as in the DEA case).

For the explanation of the observed performance it is based on a *probabilistic* formulation of the estimation problem that overcomes most limitations of previous approaches using an all in one approach or a two stage regression based approach. For more details see Daraio and Simar (2003) and Daraio (2003).

In order to control the influence of extremes values and outliers it measures the productivity performances and investigates on their explaining factors in a *robust* way also, using the order *m* efficiency measures.

Finally, it provides an *easily interpreted* graphical tool which is able to show the effect of external environmental variables on the performance of S&T systems.

In the following paragraph we present an application of the methodology described to an investigation of size effects on scientific research in the institutes of the Italian CNR.

3.2 An Illustration on the Italian National Research Council (CNR) Data

Founded in 1923, the CNR (Consiglio Nazionale delle Ricerche) is the most important national research institution in Italy, spanning many scientific and technological areas.

For this exercise we used a detailed cross-sectional database constructed by integrating several official sources on the year 1997. Further information about the database, as well as a discussion of its limitations, are reported in Bonaccorsi and Daraio (2003b, 2003c) where a theoretical and empirical analysis on size, agglomeration, and age effects in science is reported.

In the following we describe the variables and their descriptive statistics (see Tables 2.2 and 2.3) and present the graphic obtained by applying the probabilistic and robust approach described above (Figure 2.1)[7].

Table 2.2. Definition of inputs, outputs and external factors

	Variable	*Description*
Input 1	T_RES	Number of researchers
Input 2	ADTECH	Number of technicians and administrative staff
Input 3	RESFUN	Total research funds
Output	INTPUB_N	Normalised number of international publications
External factor	LABCOS	Labour costs

Table 2.3. Descriptive statistics

Variable	*Mean*	*Standard. deviation*	*Min*	*Max*	*Inter quartile range*
T_RES	13.1	9.1	1.0	45.0	11.2
ADTECH	13.8	12.8	1.0	69.0	11.0
RESFUN	984.1	865.0	45.0	7,329.0	718.0
INTPUB_N	1.0	0.6	0.03	3.1	0.8
LABCOS	2,127.4	1740.4	96.0	9,128.0	1,849.8

As explained in Daraio and Simar (2003), in order to detect the global effect of the external factor on the performance of the firms analysed, it is of interest to analyse the behaviour of the scatterplot and the smoothed regression of the ratios Q^z on Z. Q^z is the ratio between the efficiency score of a unit taking into account the external factor Z (efficiency conditional to Z) and the unconditional efficiency score. In order to have a robust measure of this effect it is reported also the robust nonparametric version plot (see bottom panel of Figure 2.1 where the plot of $Q^z{}_m$ against Z is reported). $Q^z{}_m$ is the ratio between the conditional (to Z) robust order m efficiency measures and the unconditional robust efficiency measures of a research unit analysed. We choose a level of robustness at 10% and then we find the value of m that left out the 10% of best performers in the population. In an input-oriented framework (as adopted here) an *increasing* nonparametric regression line indicates an unfavourable external factor, whilst a *decreasing* nonparametric regression line points to a favourable external factor.

[7] For a comparative productivity analysis and a bootstrap application to these data see Daraio (2003).

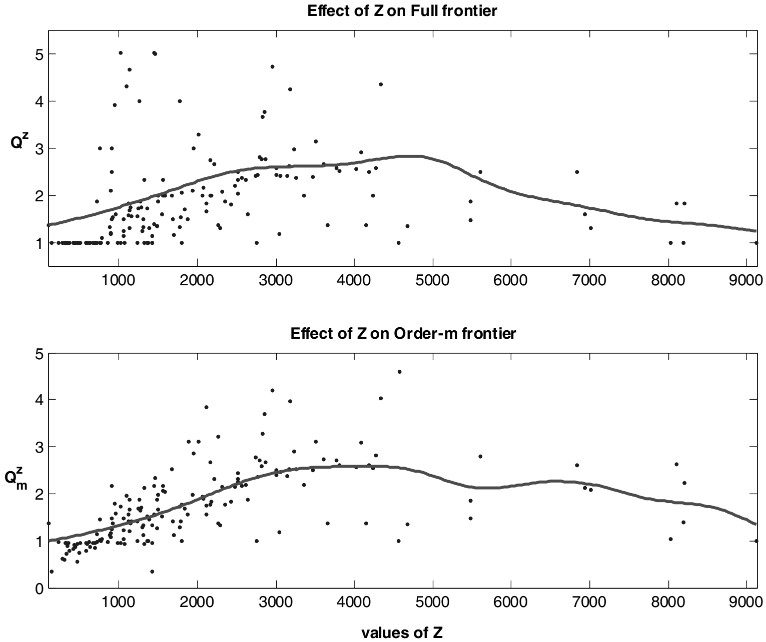


Figure 2.1. Size Effects on CNR institutes (169 obs). External factor: Labor Costs (LABCOS)

Figures 2.1 shows the effect of size (as represented by labour costs, measured in millions of Italian lire, one million Italian lire is equivalent to 516.45 Euros) on the performance of the Italian CNR institutes.

Units that lie around a Q^z value of one are not influenced by size effects, whilst units scattered in the increasing (decreasing) portion of the curve are negatively (positively) influenced by size.

A striking result is that the large majority of institutes is situated around the increasing part of the smoothed line: size *negatively* affects the performance of the majority of CNR institutes (with a level of Labour costs smaller than 4,500). Anyway, in the Italian public research system there are few large institutes (with a level of Labour costs higher than 4,500) the performance of which is *positively* affected by their large dimension. The corresponding smoothed line is decreasing, indicating a positive effect of size on their performance. This effect is confirmed if we use as proxy of size the Total Costs of research institutes (plots not reported to save place).

It is clear that these local effects could not be identified using a production function approach, in which returns to scale are summarised in a single measure. Policy implications are largely different in the two cases.

4. CONCLUSIONS

There could be some economic cases in which the function of interest can be determined by the economic theory, but one wants to reduce the strength of the assumptions required for estimation and inference. In these cases, the application of *semiparametric* statistical methods can be helpful (see Horowitz (1998) and Pagan and Ullah (1999) for an applied–oriented presentation of the several techniques available).

Nevertheless, in general situations and in complex cases the nonparametric approach seems to have several merits. In particular, in the estimation of a regression curve it presents four main advantages (Hardle, 1992). First, it provides a versatile method of exploring a general relationship between two variables. Second, it gives predictions of observations yet to be made without reference to a fixed parametric model. Third, it provides a tool for finding spurious observations by studying the influence of isolated points. Fourth, it constitutes a flexible method of substituting for missing values or interpolating between adjacent values of X.

This approach makes it possible to estimate functions of greater complexity and could be able to detect bimodal or other characteristics of distributions. The nonparametric approach is even more promising in the analysis of production frontier, particularly after the recent developments in robust techniques.

We believe that every method has some cost associated with it. Nevertheless, the diffusion and application of the developments of the econometric tools will address the main limitations.

Table 2.4 on the next page may be a useful tool, listing some basic references for researchers who wish to address empirically the difficult task of analysing productivity and efficiency in science and technology.

ACKNOWLEDGEMENTS

This Chapter has been prepared under the PRIME Thematic Network supported by the European Commission, the 6th Framework Programme and the Italian MIUR 40% Programme "Productivity and coordination in public sector research".

We gratefully acknowledge the contribution of Finn Førsund, Shawna Grosskopf, Bronwyn H. Hall, Jacques Mairesse, Camilla Mastromarco, Léopold Simar, Paula Stephan, Jerry Thursby, and Angelo Zago, who provided useful suggestions on the literature. We are the only responsible for the content of the Chapter.

Table 2.4. Econometric tools for measuring productivity: a theoretical framework and some references

	Parametric framework	*Semiparametric framework*	*Nonparametric framework*
Production functions	Griliches and Mairesse (1998) Greene (2000)	Pagan and Ullah (1999), Horowitz (1998)	Härdle (1994), Pagan and Ullah (1999)
Production frontiers	Aigner and Chu (1968), Meusen and van den Broeck (1977), Aigner, Lovell and Schmidt (1979), Kumbhakar and Lovell (2000)	Park and Simar (1994), Park, Sickles and Simar (1998, 2003)	Charnes, Cooper and Rodes (1978), Deprins, Simar and Tulkens (1984), Färe, Grosskopf and Lovell (1985, 1994), Cooper, Seiford and Tone (1999), Simar and Wilson (2003b)

REFERENCES

Abbott M., Doucouliagos C. (2003). The efficiency of Australian universities: A data envelopment analysis. *Economics of Education Review*, 22, 89–97.

Adams J, Griliches Z. (1998). Research productivity in a system of universities. *Annales d'Economie et de Statistique*, 49/50, 127–162.

Aigner, D.J., Chu S.F. (1968). On estimating the industry production function. *American Economic Review*, 58, 826–839.

Aigner, D.J., Lovell, C.A.K., Schmidt P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6, 21–37.

Allen, R., Athanassopoulos, A., Dyson, R.G., Thanassoulis, E. (1997). Weights restrictions and value judgements in data envelopment analysis: evolution, development and future directions. *Annals of Operations Research*, 73, 13–34.

Arora A., David P.A., Gambardella A. (1998). Reputation and competence in publicly funded science: estimating the effects of research group productivity. *Annales d'Economie et de Statistique*, 49/50, 163–198.

Baker B.D. (2001). Can flexible non–linear modeling tell us anything new about educational productivity? *Economics of Education Review*, 20, 81–92.

Banker, R.D., Charnes, A., Cooper W.W. (1984). Some models for estimating technical and scale inefficiencies in DEA. *Management Science*, 32, 1613–1627.

Bartelsman, E.J., Doms M. (2000). Understanding productivity: lessons from longitudinal microdata. *Journal of Economic Literature*, 38, 569–594.

Baumol W.J., Panzar J.C., Willig D.G. (1982). *Contestable markets and the theory of industry structure*. New York: Harcourt Brace Jovanovich.

Bessent, A., Bessent, W.E. (1980). Determining the comparative efficiency of schools through DEA. *Educational Administration Quarterly*, 16, 57–75.

Bessent A., Bessent W., Kennington J., Reagan B. (1982). An Application of mathematical programming to assess productivity in the Houston independent school district. *Management Science*, 28, 1355–1367.

Bonaccorsi A., Daraio C. (2003a). A robust nonparametric approach to the analysis of scientific productivity. *Research Evaluation*, 12 (1), 47–69.

Bonaccorsi A., Daraio C. (2003b). Age effects in scientific productivity. The case of the Italian National Research Council (CNR). *Scientometrics*, 58 (1), 47–88.

Bonaccorsi A., Daraio C. (2003c). Exploring size and agglomeration effects on public research productivity. Mimeo LEM, Scuola Superiore Sant'Anna, Pisa.

Brinkman P.T. (1981). Factors affecting instructional costs at major research universities. *Journal of Higher Education*, 52, 265–279.

Brinkman, P.T., Leslie L.L. (1986). Economies of scale in higher education: Sixty years of research. *The Review of Higher Education*, 10 (1), 1–28.

Cazals, C., Florens, J.–P., Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106, 1–25.

Charnes, A., Cooper, W.W., Rhodes, E.L. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429–444.

Cohn, E., Rhine, S.L.W., Santos, M.C. (1989). Institutions of higher education as multi-product firms: economies of scale and scope. *Review of Economics and Statistics*, 71 (May), 284–290.

Cooper, S.T., Cohn, E. (1997). Estimation of a frontier production function for the South Carolina educational process. *Economics of Education Review*, 16 (3), 313–327.

Cooper, W.W., Seiford, L.M., Tone, K. (1999). *Data envelopment analysis: a comprehensive text with models, applications, references and DEA-Solver software*. Kluwer Academic Publishers, Boston.

Daneshvarty, N., Clauretie, T.M. (2001). Efficiency and costs in education: year–round versus traditional schedules. *Economics of Education Review*, 20, 279–287.

Daraio, C. (2003). *Comparative efficiency and productivity analysis based on nonparametric and robust nonparametric methods. Methodology and Applications*. Doctoral dissertation, Scuola Superiore Sant'Anna, Pisa (Italy).

Daraio, C., Simar, L. (2003). Introducing environmental variables in nonparametric frontier estimation: a probabilistic approach, Discussion Paper no. 0313, Institut de Statistique, UCL, Belgium, *forthcoming* on *The Journal of Productivity Analysis*.

De Groot, H., McMahon, W.W., Volkwein, J.F. (1991). The cost structure of American research universities. *Review of Economics and Statistics*, 424–451.

Deprins, D., Simar, L., Tulkens, H. (1984). *Measuring labor efficiency in post offices*, in M. Marchand, P. Pestieau, H. Tulkens (Eds.), The performance of public enterprises – concepts and measurement (pp. 243-267). Amsterdam, North-Holland.

Dewey, J., Husted, T.A., Kenny, L.W. (2000). The ineffectiveness of school inputs: a product of misspecification? *Economics of Education Review*, 19, 27–45.

Dunbar, H., Lewis, D.R. (1995). Departmental productivity in American universities: economies of scale and scope. *Economics of Education Review*, 14, 119–144.

Efron, B., Tibshirani, R.J. (1993). *An introduction to the Bootstrap*. Chapman and Hall, NY.

Färe, R., Grosskopf, S., Lovell, C.A.K. (1985). *The measurement of efficiency of production*. Boston: Kluwer Academic Publishing.

Färe, R., Grosskopf, S., Lovell, C.A.K. (1988). An indirect approach to the evaluation of producer performance. *Journal of Public Economics*, 37, 71–89.

Färe, R., Grosskopf, S., Lovell, C.A.K. (1994). *Production Frontiers*. Cambridge University Press, Cambridge.

Färe, R., Grosskopf, S., Russell, R.R. (1998). *Index numbers: essays in honour of Sten Malmquist*. Boston: Kluwer Academic Publishers.

Färe, R., Grosskopf, S., Weber, W. (1989). Measuring school district performance. *Public Finance Quarterly*, 17, 409–428.

Farrell, M.J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, Series A, CXX, Part 3, 253–290.

Figlio, D.N. (1999). Functional form and the estimated effects of school resources. *Economics of Education Review*, 18, 241–252.

Florens, J.P., Simar, L. (2002). Parametric approximations of nonparametric frontier. Discussion Paper no. 0222, Institut de Statistique, UCL, Belgium, *forthcoming* on the *Journal of Econometrics*.

Førsund, F.R., Kalhagen, K.O. (1999). *Efficiency and productivity of Norwegian Colleges*, in Georg Westermann (ed.), Data envelopment analysis in the service sector (pp. 269–308). Wiesbaden: Deutscher Universitäts-Verlag.

Fox, M.F. (1983). Publication productivity among scientists: a critical review. *Social Studies of Science*, 13.

Greene, W.H. (2000). Econometric Analysis. Prentice Hall International, UK.

Griliches, Z., Ringstad, V. (1971). *Economies of scale and the form of the production function*. Amsterdam: North-Holland.

Griliches, Z., Mairesse, J. (1998). *Production functions : The search for identification*, in Z. Griliches (Ed.), Practising econometrics: essays in method and application, Edward Elgar (pp. 383–411). Also in S. Ström (Ed.), Econometrics and economic theory in the 20th century: the Ragnar Frish Centennial Symposium (pp. 169–203). Cambridge University Press.

Grosskopf, S., Hayes, K., Taylor, L., Weber, W. (1997). Budget–constrained frontier measures of fiscal equality and efficiency in schooling. *Review of Economics and Statistics*, 79 (1), 116–124.

Grosskopf, S., Hayes, K., et al. (1999). Anticipating the consequences of school reform: a new use of DEA, *Management Science*, 45, 608–620.

Grosskopf, S., Hayes, K., Taylor, L., Weber, W. (2001). On the determinants of school district efficiency: competition and monitoring, *Journal of Urban Economics*, 49, 453–478.

Grosskopf, S., Moutray, C. (2001). Evaluating performance in Chicago public high schools in the wake of decentralization. *Economics of Education Review*, 20, 1–14.

Gyimah–Brempong, K., Gyapong, A. (1992). Elasticities of factor substitution in the production of education. *Economics of Education Review*, 11, 205–217.

Hall, B., Mairesse, J. (1995). Exploring the relationship between R&D and productivity in French manufacturing firms. *Journal of Econometrics*, 65, 263–293.

Hall, B., Mairesse, J. (1996). *Estimating the productivity of research and development in french and United States manufacturing firms : an exploration of simultaneity issues with GMM Methods*, in K. Wagner and B. Van Ark (Eds.), International productivity differences and their explanations (pp. 285–315). Elsevier Science.

Hall, P., Simar, L. (2002). Estimating a changepoint, boundary or frontier in the presence of observation error. *Journal of the American Statistical Association*, 97, 523–534.

Halme, M., Joro, T., Korhonen, P., Salo, S., Wallenius, J. (2000). Value efficiency analysis for incorporating preference information in DEA. *Management Science*, 45, 103–115.

Hansen, L.P. (1982). Large sample properties of general method of moment estimators, *Econometrica*, 50, 1029–1054.

Hanushek, E. (1986). The economics of schooling. *Journal of Economic Literature*, 24, 1141–1177.

Hanushek, E., Rivkin, S., Taylor, L. (1996). Aggregation and the estimated effects of school resources. *Review of Economics and Statistics*, 78, 611–627.

Hardle, W. (1992). *Applied Nonparametric Regression*. Cambridge University Press.

Heathfield, D.F., Wibe, S. (1987). *An introduction to cost and production functions*. MacMillan, London.

Holbrook, J.A.D. (1992). Basic indicators of scientific and technological performance. *Science and Public Policy*, 19 (5), 267–273.

Horowitz, J.L. (1998). *Semiparametric methods in econometrics, Lecture Notes in Statistics*, Vol 131, Springer-Verlag.

Johnes, G., Johnes, J. (1993). *Measuring the research performance of UK economics departments. An application of Data Envelopment Analysis*. Oxford Economic Papers, 45, 332–347.

Johnes, J., Johnes, G. (1995). Research funding and performance in UK university departments of economics. A frontier analysis. *Economics of Education Review*, 14 (3), 301–314.

Johnston, R. (1993). Effects of resource concentration on research performance. *Higher Education*, 28.

King, W.D. (1997). Input and output substitution in higher education. *Economics Letters,* 47, 107–111.

Kneip, A., Simar, L. (1996). A general framework for frontier estimation with panel data. *The Journal of Productivity Analysis*, 7, 187–212.

Korhonen, P., Tainio, R., Wallenius J. (2001). Value efficiency analysis of academic research. *European Journal of Operational Research*, 130, 121–132.

Kumbhakar, S.C., Lovell, C.A.K. (2000). *Stochastic Frontier Analysis*. Cambridge University Press, UK.

Lloyd, P., Morgan, M., Williams, R. (1993). Amalgamations of universities: are there economies of size and scope? *Applied Economics*, 25, 1081–1092.

Mairesse, J., Sassenou, M.(1991). R&D and productivity: a survey of econometric studies at the firm level. *Science–Technology Industry Review*, 8, 9–43. Paris, OECD.

Nadiri, M.I. (1970). Some approaches to the theory and measurement of Total Factor Productivity: A survey. *Journal of Economic Literature*, 8 (4), 1137–1177.

Narin, F., Breitzman, A., (1995). Inventive productivity. *Research Policy*, 24 (4), 507–519.

Nelson, R., Hevert, K.T. (1992). Effect of class size on economies of scale and marginal costs in higher education. *Applied Economics*, 24, 473–482.

Pagan, A., Ullah, A. (1999). *Nonparametric Econometrics*. Cambridge University Press.

Park, B.U., Simar, L. (1994). Efficient semiparametric estimation in a stochastic frontier model. *Journal of the American Statistical Association*, 89 (427), 929–935.

Park, B.U. Sickles, R.C., Simar, L. (1998). Stochastic panel frontiers: A semiparametric approach. *Journal of Econometrics*, 84, 273–301.

Park, B., Sickles, R., Simar, L. (2003). Semiparametric efficient estimation of AR(1) panel data models, Discussion paper 0020, Institut de Statistique, Université Catholique de Louvain, Belgium, to appear in *Journal of Econometrics*.

Pedraja–Chaparro, R., Salinas–Jimenes, J., Smith, J., Smith, P.(1997). On the role of weight restrictions in DEA. *The Journal of Productivity Analysis*, 8, 215–230.

Pritchett, L., Filmer, D. (1999). What education production functions really show: a positive theory of education expenditures, *Economics of Education Review*, 18, 223–239.

Ramsden, P. (1994). Describing and explaining research productivity. *Higher Education*, 28.

Rizzi, D. (1999). L'efficienza dei dipartimenti dell'Università Ca' Foscari di Venezia via DEA e DFA. Nota di Lavoro 99.09, Università Ca' Foscari di Venezia.

Rousseau, S., Rousseau, R. (1997). Data envelopment analysis as a tool for constructing scientometric indicators. *Scientometrics*, 40, 45–56.

Rousseau, S., Rousseau, R. (1998). The scientific wealth of European nations: taking effectiveness into account. *Scientometrics*, 42, 75–87.

Sarrico, C.S., Hogan, S.M., Dyson, R.G., Athanassopoulos, A.D. (1997). Data envelopment analysis and university selection. *Journal of the Operational Research Society*, 48, 1163–1177.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.

Simar, L. (2003). How to improve the performance of DEA/FDH estimators in the presence of noise? Discussion Paper to appear, Institut de Statistique, UCL, Belgium.

Simar, L., Wilson, P.W. (1999). Estimating and bootstrapping Malmquist indices. *European Journal of Operational Research*, 115, 459–471.

Simar, L., Wilson, P.W. (2000). Statistical inference in nonparametric frontier models: The State of the Art. *The Journal of Productivity Analysis*, 13, 49–78.

Simar, L., Wilson, P.W. (2001). Testing restrictions in nonparametric efficiency models. *Communications in Statistics*, 30 (1), 159–184.

Simar, L., Wilson, P.W. (2002). Nonparametric tests of returns to scale. *European Journal of Operational Research*, 139, 115–132.

Simar, L., Wilson, P. (2003a). Estimation and inference in two-stage, semiparametric models of production processes. Discussion Paper no. 0307, Institut de Statistique, UCL, Belgium.

Simar, L., Wilson, P. (2003b). Efficiency analysis: the statistical approach, Lectures Notes. Institute of Statistics, UCL, Belgium.

Simar, L., Zelenyuk, V. (2003). Statistical inference for aggregates of Farrell–type efficiencies. Discussion Paper no. 0324. Institut de Statistique, UCL, Belgium.

Stephan, P.E. (1996). The economics of science. *Journal of Economic Literature*, 34, 1199-1235.

Stephan, P.E., Levin, G. (1996). The critical importance of careers in collaborative scientific research, *Revue d'économie Industrielle*, 79 (1), 45-61.

Thanassoulis, E., Dunstan, P. (1994). Guiding schools to improved performance using data envelopment analysis: an illustration with data from a local education authority. *Journal of the Operational Research Society*, 45 (11), 1247–1262.

Thursby, J. (2000). What do we say about ourselves and what does it mean? Yet another look at economics department research. *Journal of Economic Literature*, 38 (2), 383–404.

Thursby, J., Kemp, S. (2002). Growth and productive efficiency of university intellectual property licensing. *Research Policy*, 31, 109–124.

Thursby, J., Thursby, M. (2002). Who is selling the ivory tower? Sources of growth in university licensing. *Management Science*, 48(1), 90–104.

Verry D.W., Layard, P.R. (1975). Cost functions for university teaching and research, *Economic Journal*, 85, 55–74.

Verry, D.W., Davis, B. (1976). University costs and outputs. Amsterdam: Elsevier.

Zhang, Y., Bartels, R. (1998). The effect of sample size on mean efficiency in DEA with application to electricity distribution in Australia, Sweden and New Zealand. *The Journal of Productivity Analysis*, 9, 187–204.

Zhu, J. (1996). Data envelopment analysis with preference structure. *Journal of the Operational Research Society*, 47, 136–150.

Chapter 3

INDICATORS FOR NATIONAL SCIENCE AND TECHNOLOGY POLICY

Their Development, Use, and Possible Misuse

Hariolf Grupp[1] and Mary Ellen Mogee[2]

[1]*Karlsruhe University and Fraunhofer ISI, Karlsruhe, Germany*
E-mail: h.grupp@isi.fraunhofer.de
[2]*Mogee Research & Analysis, Reston, Virginia, USA*

Abstract: The purpose of this chapter is to present a survey of the development of Science and Technology (S&T) indicators and their use in national policy making as well as to provide evidence of the vulnerability of S&T indicators to manipulation. A brief history of the development of S&T indicators begins with the United States followed by their worldwide diffusion, with particular emphasis on Europe. The current status of S&T indicators and newer developments towards composite indicators, benchmarking, and scoreboarding is discussed. To investigate the robustness of innovation scoreboards empirically a sensitivity analysis of one selected case is presented. It is shown that composite scores and rank positions can vary considerably, depending on the selection process. It seems not to be too difficult to argue for a ‘country friendly’ selection and corresponding weighting of indicators. Thus the use of scoreboards opens space for manipulation in the policymaking system. Further research is needed on alternative methods of calculation to prevent their misuse and abuse.

1. INTRODUCTION

The purpose of this chapter is to present a survey of the development of science and technology (S&T) indicators and their use in national policy making, to provide evidence of the vulnerability of S&T indicators to

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 75-94.

manipulation and to suggest questions for future research.[1] The chapter is organised as follows. In the second section we review briefly the history and development of S&T indicators and the major S&T indicators reports in the United States, Japan, and Europe. The third section presents a sensitivity analysis of a composite indicator. Section 4 presents our conclusions and suggestions for future research.

2. A BRIEF HISTORY OF THE DEVELOPMENT OF S&T INDICATORS

2.1 The Beginning — the United States

Science and technology indicators had their start in the United States. The first *Science Indicators* report was published in 1973. The National Science Board (NSB), the policymaking board of the National Science Foundation (NSF), was mandated by the US Congress to publish the report biennially. (Perhaps not surprisingly, given the NSF's pre-eminent support for basic scientific research, it was not until 1987 that the focus and title of the report were broadened to *Science and Engineering Indicators*). In practice the report has been prepared by the Science and Engineering Indicators Unit (SIU) in the NSF's Science Resources Studies division and reviewed by the NSB, which also prepares a brief discussion piece as part of the report.

The idea of 'science indicators' (SI) was an outgrowth of the move toward 'social indicators', i.e., indicators similar to economic indicators such as Gross National Product (GNP) that would provide measures of the health of society. Based on the model of economic indicators, some observers expected science indicators to be a narrow range of statistics that tells about a larger universe, in the same way that new housing starts tell how the economy in general is doing. To date the science indicators report has not evolved in this direction, but rather has become a compendium of many different statistics to measure the health of US science and technology and to compare the US with other nations.

In the early years there was considerable criticism of the science indicators report including the publication– and patent–based indicators. The NSB member and famous mathematician Saunders MacLane criticised the

[1] An earlier, shorter version of this chapter was presented at the Conference in honour of Keith Pavitt, University of Sussex, November 13th–15th, 2003, and subsequently published in a partly overlapping way in a Special Issue of the journal 'Research Policy'.

publication–based indicators, in particular. Early SI reports counted publications by country instead of by individual's addresses or institution, which resulted in the finding that one third were British. In response to the critical writing the NSF held a world conference on the coverage and validity of the set of journals and the way of counting publications.

Similarly there was considerable scepticism in the beginning about patents as science or technology indicators. The NSF had to do lots of studies to show that the use of patents had methodological backing. However, it was patent indicators which showed the US that it should pay more attention to Japan as a competitive industrial power. In the part devoted to international comparisons increases were found in Japanese R&D funding and in all categories of patenting. This finding — that Japan was a power to contend with — was surprising to many at the time.

In addition to criticism of data and methodologies used in the SI report, some criticism of science indicators was based on resistance in the science community to making government funding decisions based on quantitative indicators. Many scientists believed, and continue to believe, that such decisions should be made on the basis of peer review.

In 1985 the House of Representatives Committee on Science and Technology undertook a Science Policy Study which asked the Office of Technology Assessment to examine "… the extent to which decision making would be improved through the use of quantitative mechanisms associated with the concept of investment". OTA concluded that "… while there are some quantitative techniques that may be of use to Congress in evaluating specific areas of research, basic science is not amenable to the type of economic analysis that might be used for applied research or product development". In his accompanying letter, John H. Gibbons, the OTA Director, stated further:

> "Much of the vitality of the American research system lies in its complex and pluralistic nature. Scientists, citizens, administrators, and Members of Congress all play various roles leading to final decisions on funding. While there may be ways to improve the overall process, reliance on economic quantitative methods is not promising. Expert analysis, openness, experience, and considered judgement are better tools."

Despite resistance, over time wide acceptance of the indicators grew. This has been owed in part to pressure from users and decision makers who wanted to be able to show that investment in S&T had value and impact. In response, the SI unit moved beyond the original indicators to additional indicators of interest to policy makers. There was constant improvement by taking feedback and criticism into account.

It was recognised early on that some indicators represent science and engineering resources or 'inputs' to the process of science and engineering, whilst others represented the process itself ('throughput' or 'flow'), the results or 'outputs' of science and engineering, and the effects or 'impacts' of science engineering. The early indicators tended to be heavy on inputs and throughput and weaker on outputs and impacts, spurring the SIU to look for and develop more indicators of outputs and impacts.

In addition to indicators of the level of scientific and technological activity, indicators were developed which examined international collaboration and intersectoral — e.g., university–industry — collaboration. Funding, citations, and co-authorship were all used as indicators of collaboration.

The SIU also developed indicators of the importance of science and engineering. For instance they looked at publication and patent. They also looked at patents citing literature (see, e.g. Narin and Noma, 1985). Publication citations have been used to indicate the quality of research. Patent citations have been used variously to indicate 'technological significance', 'social economic value' and 'private economic value' (for the many applications of patent analysis, see Breitzmann and Mogee, 2002). Other measures of value or quality from patent indicators include the rate at which patent maintenance fees are paid and the number of patent countries in which patent protection is sought for an invention (patent family size).

Later under the Government Performance and Results Act (GPRA) and the pressure to show use of government funded basic research, one of the key indicators that was helpful to policy makers was the patent citations to basic research. This was used to show that commercial technology was building on the basic research funded by government. Other agencies doing this today are the Office of Naval Research (ONR), the Department of Defense (DOD), the Department of Energy (DoE) and the Department of Veterans Affairs (DVA). Industry also uses patent citations and their references to basic research. It is difficult for funding agencies to show a direct use for the basic research they support. The Office of Management and Budget (OMB) and the Congress both review GPRA reports, and the OMB, in particular, likes numbers, so research funding agencies have used these numbers on patents citing basic research to argue for higher budgets.

The SI unit conducted a series of 4–5 surveys of the usage of the SI report. These were generally biennial, like the Science & Engineering report, and were sent out with the report. At first the questions were not split up by what specific indicators the people were using. Instead they asked about the policy issues the respondents wanted to cover and what they found useful. The users always asked for output indicators. Publications and patents are both output indicators. Therefore they had high priority for SI and they still do. Responses were received from the NSB, scientific community, the Office of Science and Technology Policy (OSTP), OMB, and the OECD, amongst others. They all wanted the indicators to be relevant and useful.

The National Science Board (NSB) and the Director of the National Science Foundation (NSF) are important users of the publication data in the SI report. This is natural because publication indicators pertain largely to academic research, and the NSB and the NSF Director are major players in US science policy. This kind of information is reportedly used in allocating funds between scientific fields. Similarly, when they first found evidence of the high citation rate from patents to basic research, the NSB used it to make the industrial use of basic research the centerpiece of their discussion piece in the SI report.

That the NSB has continued the publication of the SI report, although it make suggestions for changes, suggests that the members find the compilation to be useful. In response to a recent (2003) proposal to cut back the SI report and to include just those statistics that are not published elsewhere, the NSB member Anita Jones stated in a presentation that:

– *Science & Engineering Indicators* (S&EI) is a leading data source for R&D policymakers;
– The data are sound; the definitions and categories change slowly;
– The longitudinal data emphasis is very useful;
– The data are collected in one place and repeatedly updated.

Dr. Jones went on to say that the audience for S&EI includes federal and state political appointees in the R&D area, scientists and engineers on advisory boards such as the NSB, the NSF directorate advisory committee, many levels of agency advisory committees, and National Academies' task forces – all important players in the US science policy system. In Dr. Jones' words, "S&EI is the 'one stop shop' for policy makers who do not study the multiplicity of R&D statistics publications." (July 2003)

The authors are not aware of any instrumental use of publication- or patent–based indicators in national policy, that is, cases in which a decision hinged upon the publication or patent indicators. There is wide recognition, however, that statistics and indicators are often used to justify decisions or to support a particular side in a disputed issue. It is also recognised that there is not necessarily anything wrong with using statistics in this manner, but the limitations and meaning of the data need to be made clear when they *are* used in this way. These types of use of S&T indicators present substantial potential for abuse.

Today the indicators in the S&EI report are in widespread use in the US. A broad range of participants in the policy process cites the statistics. Particularly since the federal government came under pressure from the

GPRA, S&T indicators have been reinvigorated in their use. They tend to be used to show where the US stands with respect to other countries and as a reason for increasing funding to particular areas of science or technology.

Also in the US, the Council on Competitiveness has developed its own set of innovation indicators. It uses patent data together with industry analysis. They are thinking about repeating their innovation summit. It is a lot of work and financially difficult in both publications and patents (particularly patents). It needs to be done in a collaborative way.

2.2 Worldwide Diffusion of S&T Indicators and the OECD

Today many other countries use the US data or have been inspired by the US to develop their own indicators and indicator systems. The US is still encouraging the spread of science indicators by consulting with countries which are establishing systems to track and use science indicators. The indicators are broadly accepted. For example, Latin America has become involved in S&T indicators. SI helped RICYT (La Red Iberoamericana de Indicadores de Ciencia y Tecnología) to start a S&T indicators network. Over the past 5–6 years they have developed comparable data such as the SEI pocketbook data book. It covers all the Americas, including the US and Canada, plus Spain and Portugal. It includes a broad range of indicators such as R&D funding, publications, and patents.

It may be an overstatement to say that other countries have gone further than the US in the use of indicators. However, it seems that European countries may be going in different directions in the use of S&T indicators, using indicators more in benchmarking and in foresight exercises than the US (see section 2.3 below).

The idea of developing a S&T indicator system that includes publications and patents and its use by policy makers is taken further in certain countries and used more rigidly in terms of funding individual researchers than it is in the US. Countries which have a more centralised science policy system tend to use science and technology indicators in a more rigorous way. For example, in France and Mexico quantitative indicators are used in the decision to give individual researchers more funding. In the US S&T indicators are used more widely by management for general policy and awareness of trends than at the level of allocating resources to individual researchers.

Beginning in the late 1970s the Organisation for Economic Co-operation and Development's (OECD) secretariat for science and technology indicators (restructured and renamed several times in the past decades) exerted a very important standardising role within the member states of the

OECD. By inviting researchers, statisticians, and other responsible persons to join workshops, by editing manuals on R&D, patent and innovation measurement, and revising unclear national statistics to OECD standards, the secretariat carried out an important task by making national scoreboards on S&T comparable. The OECD bodies resisted producing only simple aggregated rank tables of countries' innovation performance. Even the most recent STI scoreboard (OECD, 2003), the sixth in a biennial series which started a decade ago, did not produce scalar measures of innovation activities of countries, although it did give particular attention to offering new or improved official measures for international comparisons in the major areas of policy interests. This stands in clear distinction to what is observed at the European level, which will be reported on below.

2.3 S&T Indicators in Japan

In Japan, for reasons of language, major English reports on S&T and the respective indicators began with the establishment of NISTEP (National Institute of Science and Technology Policy) in the year 1988. Certainly, important Japanese sources on R&D expenditures were published before, such as the Report on the Survey of Research and Development (Management and Coordination Agency, various years) issued annually in Japanese with English sub-titles to table and figure captions. Yet the special dedication of the foundation of NISTEP was to bring Japanese S&T information, including indicators, to an international audience. Within the mission of NISTEP to contribute to policy making by taking a sort of task force, major internationally comparable reports on S&T indicators were published. These indicators systematically organised the knowledge about scientific and technological activities of Japan and the corresponding reports were the first to make the overall state of these activities quantitatively comprehensible.

In the context of this chapter, in focussing on aggregation methods of S&T indicators, the approach of NISTEP to establish a Japanese science and technology system must be recognised. The basic method of 'integrating' S&T indicators by the Japanese institution was in using a 'cascade model' and factor analysis (Niva and Tomizawa, 1995; Kodama, 1987). The international comparison of overall strengths in science and technology was processed in such a way that 13 indicators for Japan and other countries were used to illustrate national S&T activities such as inputs in R&D, staff, output, number of scientific paper citations, and so on. That is, the multiplicity of indicators was reduced in a way that looks for similarities in the structure of the data and results in a lower-dimensional array of indicators, which is more than a simple ranking. To the best of our knowledge these activities were not fully continued in the past years.

2.4 S&T Indicators in Europe

In Europe reporting on national science and technology performance has changed markedly in the past ten years. There are two main reasons for this. First, the former communist countries did not keep with OECD conventions, and the little information which was available before around 1990 was often not comparable. It was also widely considered to be systematically overestimated. This started to change around 1990 and since then several national reports from Eastern European countries have been issued, some of them in the English language. Yet in this short chapter it is impossible to give a full account in this respect (see, for instance, Gokhberg et al., 1999; CSRS, 1998).

The second change was the more active European Commission. A landmark in this respect is the first European report on S&T indicators (European Commission, 1994), which consisted of a massive attempt to collect available data of various kinds. The Commission was assisted by a large group of leading European researchers in that area. In 2003 the third such report was issued. This new role of the European Commission triggered numerous competing activities which led some observers to note an 'oversupply' of S&T indicator reports (see below).

Before these two trends made themselves felt in the 1990s and in the first years of the 21st century, a variety of non-comparable reporting systems in major Western European countries was in place. Here again we do not attempt to give a full account of the 1970s and the 1980s. Nor is it possible to do justice to every country. Some of these reports were not published regularly but only in exceptional cases and in different formats and many of them are in national languages. These activities were not stopped when the European Union level came up with own products and most of the national series continue today.

To give a few examples let us mention the French report on 'Science & Technologie – Indicateurs' which has been published since the inception of the Observatoire des Sciences et de Technologie (OST) in Paris in 1990 (OST, various years). The series of reports is clearly subdivided into the national level, the European level, and the international level. Observers within France are proud of the many data series and the systematic and continuous way the report is published. However, the reports are known to lack analytic and policy sections and assessments to complement the data series.

In the United Kingdom various related publications exist, but periodical reports in a consistent format that provide comparable information over a longer interval were not established. Amongst the newer publications let us just mention the Department of Trade and Industry (DTI) economics paper No. 7, 'Competing in the global economy – the innovation challenge' (2003).

In Germany, since 1965 the 'Bundesbericht Forschung' (Federal Research Report) is published every four years in the German language (BMBF, various years). From time to time more or less abridged English versions are available. The latest such report appeared very recently (May 2004). Because of the federal structure in Germany, this report has a national part, a state (Länder) part, and also some international comparisons. It is more focussed on R&D inputs and R&D infrastructure, and describes large organisations in Germany. This report is assisted by the 'Bericht zur technologischen Leistungsfähigkeit' (report on technological competitiveness), which has been published annually since 1985 (with various editors; for the latest version see Grupp et al., 2003). The latter report is not as complete as the former in terms of compiling R&D data and it has a less official character. It is prepared by research institutes for the German government and is quite analytic and policy oriented. Government officials occasionally were not happy with the assessments and findings. In the case of Germany one can also demonstrate the problems with former communist countries. The very complete R&D statistics of East Germany (former German Democratic Republic) had to be adjusted in a very complicated way in order to be comparable to the Western system. This was done in the first years of unification and now comparable backward information is available. This case may be taken as typical for all the Eastern European countries.

National reports are available from Austria (Pohn-Weidinger et al., 2001; Republik Österreich, 2003), Italy, the Netherlands, the Scandinavian countries, and so forth. Most of them are in the national languages.

Returning to the European Union level, in the past several years, in addition to the three European reports of S&T indicators mentioned previously, a variety of other reporting systems were established. Benchmarking activities were started with the explicit aim of going beyond existing statistics and providing new types of data not available so far (for instance, R&D staff by gender; European Commission, 2002). These activities are co-ordinated by the Directorate General for Research and assisted by a High Level Group of Experts on Benchmarking, Excellence, Co-Ordination of National Policies. The Directorate General for Research also issued the booklet 'Key Figures 2003–2004' (2003b). A preliminary version of an 'Innovation Scoreboard' was published in 2000 by the innovation/Small and Medium Sized Industry programme of the Enterprise Directorate General. Since then the European Innovation Scoreboard has been published regularly. The same directorate also published a Biotechnology Innovation Scoreboard (2003). Both types of reports make use of 'composite indicators', which we discuss in more detail below. Composite indicators are also part of chapter 1 in the most recent European Report on Science and Technology indicators (2003).

It seems that the European Commission is driving S&T indicators in the direction of aggregation of different types of indicators into simpler constructs in order to summarise complex multi-dimensional phenomena.

2.5 Current Status of S&T Indicators

To summarise their development, S&T indicators have evolved over the past 30 years to become a large number of statistics each of which describes a portion of the science and technology system. Because they are by definition partial they must be used in combination with each other and with other kinds of data such as expert opinion to provide a full picture. Every indicator has its own unique strengths and weaknesses, and indicators need to be selected according to the problem or question being addressed (Grupp, 1998).

Progress has been made toward linking particular indicators to particular parts of the S&T system or the innovation process. Researchers have moved from the simple concepts of inputs and outputs to concepts of inputs, throughputs, outputs, and impacts at *various stages of the process*. Publications can be used to indicate the output of basic scientific research, for example, but would be misleading if used (alone) to indicate the output of industrial research and technology development. Patents, on the other hand, are useful as an indicator of applied research and technology development (*loc. cit.*).

It is increasingly recognised that some indicators are appropriate in certain contexts and not in others. For example, if the objective is to understand the development of computer software in the last quarter of the twentieth century, a patent analysis would not be recommended. This is because, although software is increasingly patented, particularly in its early years it was not, so an analysis of software patents would miss much software activity in the early years. However, if one is interested in the extent to which ownership of software patents is concentrated in a few companies, an analysis of patents would make sense.

Progress has also been made in developing indicators of quality, importance, or value, although these concepts themselves have not been defined as well as they should be.

Given the above situation, S&T indicators can be misused or abused, as well as used for positive purposes (Pavitt, 1988). Some of the possible misuses include:

– Reliance on a single indicator;
– Use of an indicator that is inappropriate for the technology, system, or stage of the R&D process;
– Drawing conclusions which are too strong, given the 'indicative' nature of indicators;
– Making inferences that are inappropriate, based on the indicator and its relationship to the phenomenon of interest.

2.6 The Development of Composite Indicators and Related Concepts

To sum up several decades of debate, the measurement of science and technology requires measurements along many dimensions. To date no ideal 'catch all' variable for science or innovation has been developed (Patel & Pavitt, 1995). Therefore in many cases multiple indicators have been used. However, the use of multiple indicators means that conventional methods such as the knowledge production function (Griliches, 1995) and many other concepts of efficiency measurement cannot be supported. Optimal configurations of measurement must be worked out in some other way (such as factor or data envelopment analysis). The recognition of the need to measure multiple dimensions of science and technology has also led to the emerging and pioneering field of composite indicators to enlighten national S&T policies.

A fortiori, the multi-dimensional science and technology (S&T) variables are usually not expressed in monetary terms but rather are measured in other units (such as patent counts, innovation counts, number of citations, etc.) and may not be comparable to each other. Lacking a well defined correspondence between relevant S&T data — for instance a conversion relationship between dollars and patent numbers — the multi-dimensional profiles cannot be aggregated into an overall scalar figure. This situation is fundamentally different from one in which all variables are fully specified in terms of quantities and costs or prices such as such well known economic indicators as the Gross Domestic Product (GDP).

On the micro-level of companies or single innovation projects, decision–oriented measurement practices such as 'benchmarking' or 'scoreboarding' have become well established. Benchmarking is the practice of identifying the organisation (e.g., competitor firm) which is the best at a particular function or activity, such as innovation, and using that organisation's metrics as the goal to be achieved and surpassed. "Benchmarking is the continuous process of measuring products, services, and practices against the toughest competitors or those companies recognised as industry leaders." (Kearnes, 1986). Although this is clearly a quantitative approach, it also has qualitative aspects. Camp, for example, refers to benchmarking as "the search for industry best practises that lead to superior performance" (Camp, 1989, p. 12). Benchmarking in the industrial context is very action oriented, aimed at improving business operations and competitiveness. The development of indicators and their aggregation are the means to an end in industrial benchmarking.

Let us note here, that, before a decision is made in a firm, the available database can be questioned and even partly laid aside. Here the assessment of quality, although it may be difficult, can be solved in some way. The database with all its biases does not automatically determine decisions. This may be different in the context of a national policy (see below).

Another concept related to S&T indicators that has developed in the business world is that of scoreboarding. Like the sports scoreboards which show how many goals, runs, or points have been scored in a competition or match, scoreboards have been developed and used to show companies how they stand with respect to their competitors on aggregate, widely recognized metrics of business performance such as productivity. In the area of industrial research and development (R&D) and innovation, 'R&D Scoreboards' and even 'Patent Scoreboards' have been developed and published.

Recent years have witnessed the increasing application of these methods, relatively uncritically, for national or regional science and technology policy. In particular, the use of composite indicators is being promoted as an emerging and pioneering field (European Commission 2003 and further references given there on p. 433, in footnote 1). At this level innovation scoreboards and the like are not usually used instrumentally to make policy decisions, because decision making in science and technology policy is quite a complicated negotiation procedure between societal interests and interest groups (Edler et al., 2003). Scoreboards of national innovation performance instead function more as 'soccer league tables' telling the public which countries are performing well or second rate, which have caught up or fallen behind.

The problems with this use of benchmarks or scoreboards on a national level lie in the lack of clear theoretical models that tell us which indicators to select, how to weight them and how to handle cross-country differences in the availability of data (Pohn-Weidinger et al., 2001; European Commission, 2003a). To say the least, this use of scoreboards or benchmarking rank tables may be dangerous because the numbers provided are taken at face value with little discussion of their validity. Substantial space exists for manipulation by selection, weighting and aggregating indicators. This chapter attempts to raise this point and to provide empirical examples of the range of interpretation or misinterpretation of national innovation scoreboards.

Successful scoreboard–based analysis should depend on mastering the art of indicator selection and scoreboard design. As a *sine qua non* for reasons of public accountability, scoreboards — as any advanced evaluation method — need a clear and transparent structure and recognised concepts (Tijssen, 2003).

3. CALCULATING COMPOSITE INDICES: ONE EXAMPLE

3.1 Methodology

In this chapter we want to investigate the robustness of innovation scoreboards empirically by sensitivity analysis of one selected case. As we have argued above, this seems to be the European speciality driven to a large part by bodies of the European Commission. In any case, it seems to be a newer development within the long standing tradition of S&T indicators. “By aggregating a number of different variables, composite indicators are able to summarise the big picture in relation to a complex issue with many dimensions.” (European Commission, 2003, p. 433). What can we learn from aggregating that goes beyond the detailed information?

The procedure will be as follows: We take the original composite indicator of the European Innovation Scoreboard (2001) and compare the ranking of countries by various ranking methods, namely by

– original, Olympic, average and weighted ranks;[2]
– metric scales (weighted and un-weighted) in distinction to rank positions; and
– selective omission of data in order to ‘promote’ a countries’ position.

We determine the weights by grouping of the various indicators into input, throughput, and output and give each group one third weight, but equal weights within the group. There are many metric scales possible; in our case we use the technometric scale originating from evolutionary

[2] The original ranking is explained below, by ‘Olympic’ we mean that the ranking is done by descending numbers of ‘gold medals’ first, then by ‘silver medals’ (rank position 2), and so on.

economics. It adjusts the interval to ‘real market’ competitive positions (Grupp, 1998).

The empirical base of our case study is the European Innovation Scoreboard 2001 (see section 2.4). In that base, a combination of 18 indicators is presented, namely S&E graduates, tertiary education, lifelong learning, employment in manufacturing and services, R&D intensity, business expenditures on R&D, European and US patents, SMEs’ innovation and co-operation, innovation intensity, venture capital, new capital, new products, internet access, information technology markets, and the high tech value added.

From this scoreboard, a ‘tentative summary innovation index’ (SSI) is constructed, placing, for instance, Sweden in rank position 1 (score 6.5), the UK in rank position 4 (score 4.4), Germany in average rank position 9 (score 0.6) and Greece in rank position 16 (score -7.9) (see Figure 3.1). The SII is equal to the number of indicators which are 20 per cent above average minus the number of indicators that are 20 per cent below. The index is normalised to the interval [10, –10]. An index of zero represents the EU average.

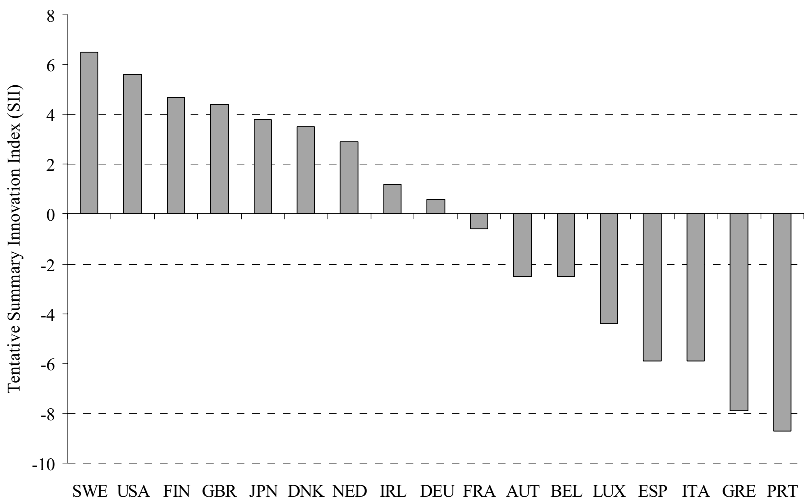


Figure 3.1. Summary innovation index of the European Innovation Scoreboard 2001 (original graph from p. 12)

One may ask several questions, for instance, why this selection of indicators and why is the aggregation done in such a peculiar way? Is it justifiable to give equal weights to 16 out of the 18 indicators but count the two sets of patent data by half each? In addition internal criticism was raised: “While this technique could prove suitable when we have the same number of indicators across countries, its relevance declines sharply when observations are unevenly distributed across countries, as is the case here." (European Trend Chart, 2003, p. 16, footnote 7). In fact, we face the problem that some data for some countries are missing, which raises special problems. Because we consider benchmarking and score boarding designs an art, we do not want to continue with arguments whether or not precisely this procedure is the best solution of all possible alternatives. But we rather want to process the given data in some other ways and compare the sensitivity of the results (the country ranking) to the original method.

3.2 Robustness of Composite Indices — Selected Results of a Sensitivity Analysis

In Figure 3.2 we provide the results of several aggregation procedures for composite indicators other than the original, as suggested in section 5.[3] What we learn from Figure 3.2 is that Sweden is in rank position 1 irrespective of the aggregation procedure. The same is true for Finland in rank position 2. All the other countries vary by one to four rank positions depending on the aggregation procedure, but overall the impression is that the country ranking cannot be completely turned upside down. The countries in top positions are always in a good position and those at the end of the scale are not positioned in the first half of the league irrespective of the aggregation details.

In particular, the average ranking and the weighted average ranking are most similar to the original SSI index, whereas the Olympic scale can change the picture more seriously. Consider the case of France: France is nowhere in best position amongst the 18 variables, thus wins now a 'gold medal' and will be placed behind all other countries with at least one gold medal (number one in one of the 18 variables).

[3] For EU countries only, thus omitting the US and Japan.

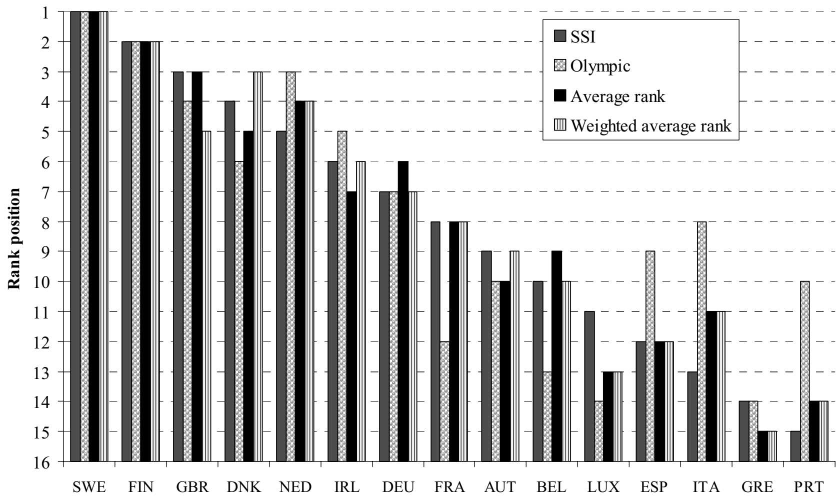


Figure 3.2. Results by selected ranking procedures

Also interesting is the case of the United Kingdom. The UK performs quite strongly in most input indicators but less so in throughput and output variables. As the set of the 18 indicators is very much input biased, equal weighing puts the United Kingdom in a favourable position. When one starts to give all inputs together the same weight as all throughputs and all outputs then this 'natural advantage' in the original SSI ranking vanishes for this country and other countries catch up.

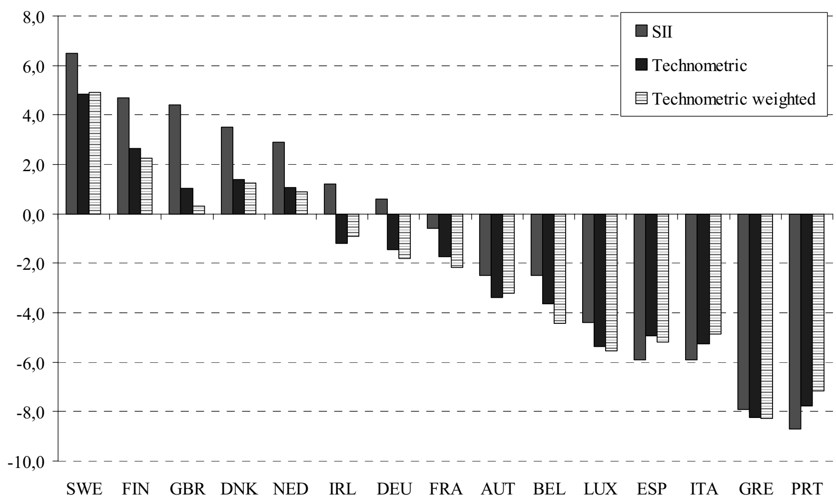


Figure 3.3. Results by selected metric procedures

Figure 3.3 displays the results by comparing the original SII index with metric procedures. Rank tables bear the problem that the distance between any two adjacent positions can be very small in original indicator values or can be large. Metric scales, in distinction, conserve the distances of the original variable values and transform them similarly.

This procedure yields quite different rankings, as is shown in Figure 3.3. Take the example of the United Kingdom again. In the SII scoreboard the country is in third place within the EU countries); however, the distance in most variables to the leading countries is much larger than the rank positions suggests; the United Kingdom seems to be closer to European average if weighted metrics scales are used. Again, the difference is still more pronounced if the weighted metrics are taken, because the 'natural advantage' in input variables vanishes. In the case of Germany the SII index is just above EU average whereas the metric values are clearly below.

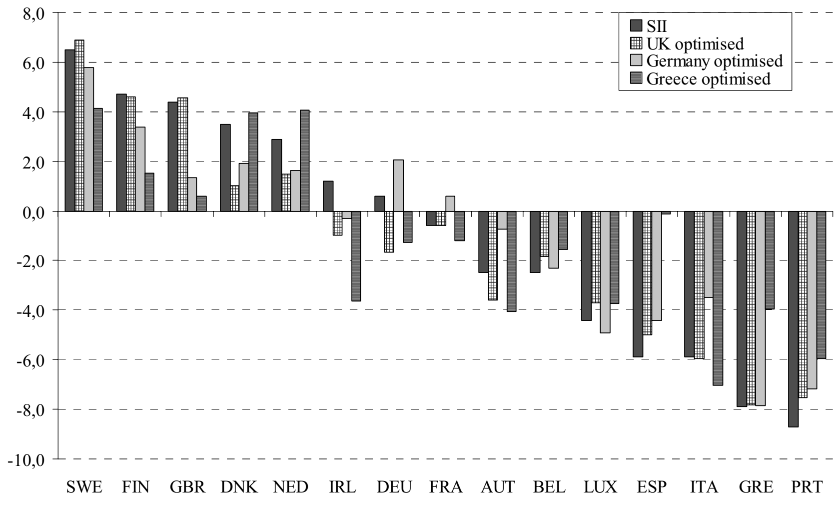


Figure 3.4. Some results by selective omission of variables

A third type of sensitivity exercise consists in selective omission of certain variables. We already have argued that scoreboard design is more an art than a science and it is difficult to argue why 18 indicators are used, not 16 or 20, and why exactly these. If we now tune the calculation by selective omission of those variables in which a certain country does not perform very well, we can try to tune the selection of variables in favour of some countries. In Figure 3.4 we present the results of an optimisation for the UK, for Germany, and for Greece.

The optimisation attempt for the UK is not very successful. The reason is that the set of 18 variables is already optimised in favour of the UK by the original SII (with its many input variables). If we exclude some output and throughput variables in which the UK does not perform very well we cannot really improve the countries' position compared to the original index. This is true for all countries being in lead positions in the original scale; these benefit from the selection by the original scoreboard and cannot really be pushed ahead further.

For countries in middle places and further down the scale one can optimise their position with more success. For instance, Germany can be put in rank position 3 by optimising the selection of indicators, because Germany performs mediocre in life long learning, venture capital, and other variables. If these are taken out of the indicator set the country's position improves considerably. The same is true for Greece, which can be brought upwards by several rank positions if a selection of variables being favourable towards Greece's performance is taken.

How plausible is country tuning? The optimised UK index is achieved by not considering EU patents and giving more weight to US data, not considering business expenditures on R&D but giving more weight to gross expenditures on R&D and to venture capital. These assumptions are not really revolutionary; in particular, in the case of two patent data sets it is questionable why both of them should go into the summary index with half weight (as is the case for the SII). Germany's index, as has been mentioned above, may be improved by not considering venture capital, the opposite assumption as for the UK – but is venture capital really a traditional core S&T indicator? Greece's profile profits from not considering patent indicators at all and leaving out high tech value added.

All these assumptions seem to be soft and can certainly be discussed seriously. Altogether, we think the selection of any one set of indicators does not give equal justice to all countries and thus the selection problem is implicitly a way to tune country positions. This is done by disputable arguments and does not need any heroic assumptions.

4. DISCUSSION AND SUGGESTED RESEARCH QUESTIONS

In this chapter we have given a very brief survey of reports on science and technology indicators in the triad regions. We argue that the European Commission is a latecomer, but is playing a more active role in recent years and is driven in the direction of composite indicators, which do not seem to be of primary concern in most single country S&T indicator efforts.

By applying alternative aggregation procedures to one selected example (introduced by EU bodies) different weights and a different selection of indicators, we have shown that summary scores and rank positions can vary considerably. It seems that ranking is less sensitive to calculating procedures but gives no information on the size of the gaps. Metric scales seem to provide more insights into relative positions of nations in the very sense of benchmarking. Sensitivity analysis further shows that exclusion or inclusion of variables tend to be a bigger problem than a slight variation of indicator scales. It seems not to be too difficult to argue for a 'country friendly' selection and corresponding weighing of indicators. To say the least, this use of scoreboards or benchmarking tables may be dangerous if the summary numbers provided are taken as such with little discussion of their validity.

The space for manipulation of scoreboards by selection, weighing and aggregation is great. Further research should remedy the situation. This chapter attempts to raise this point and to provide empirical examples of the range of interpretation or misinterpretation of national innovation scoreboards. It did not attempt to suggest more viable alternatives.

Triggering the discussion of these problems, an alternative to scoreboarding of national innovation indices is, nevertheless, suggesting, namely, the use of interval based metric scales in order not to hide the size of the gaps. The use of multi-dimensional representations is the minimum requirement, such as 'spider' charts. Maps of similarity between country structures in science and technology may have more explanatory power in particular when combined with non-quantitative methods.

More research is needed on the validity of S&T indicators, their relationship to important S&T policy concepts, their performance in different science and technology domains, their sensitivity to selection, inclusion, and alternative methods of calculation, as well as their use in the policymaking system and means of preventing their misuse and abuse.

REFERENCES

BMBF (Ed.). (various years). Bundesbericht Forschung. Bonn/Berlin.

Breitzman, A.F., Mogee, M.E. (2002). The many applications of patent analysis. Journal of Information Science, 28 (3), 187–205.

Camp, R.C. (1989). Benchmarking: the search for industry best practices that lead to superior performance. Milwaukee: ASQC Quality Press.

CSRS (Ed.). (1998). Technological innovation in Russia. Moscow: Centre for Science Research and Statistics.

DTI (Ed.). (2003). Competing in the global economy.The innovation challenge. London: DTI.

Edler, J., Kuhlmann, S., Behrens, M. (Eds.). (2003). Changing governance of research and technology policy – The European research area. Cheltenham: Edward Elgar Publishing.

European Commission (Ed.). (1994). The European report on science & technology indicators 1994. Brussels: EUR 15897 EN.

European Commission (Ed.). (2001). European innovation scoreboard 2001. Luxemburg: SEC (2001), 1414.

European Commission (Ed.). (2003). Third European report on science & technology indicators 2003. Brussels: EUR 20025 EN.

European Commission (Ed.). (2003a). Biotechnology innovation scoreboard 2003. Brussels: Innovation/SMEs Programme.

European Commission (Ed.). (2002). Benchmarking national research policies. Brussels: EUR 20494 EN.

European Commission (Ed.). (2003b). Towards a European research area science, technology and innovation – key figures 2003–2004. Brussels: EUR 20735 EN.

Gokhberg, L.N. Gorodnikova and J. Wuttke (1999). Volkswirtschaft im Übergang – Ein Vergleich der FuE-Indikatoren in Russland und Deutschland. Moscow and Essen: CSRS and Stifterverband.

Griliches, Z. (1995). R&D and productivity: econometric results and measurement issues. In Stoneman (Ed.), (pp. 52–89).

Grupp H., Legler, H., Gehrke, B., Breitschopf, B. (2003). Zur technologischen Leistungsfähigkeit Deutschlands 2002. Berlin: BMBF.

Grupp, H. (1998). Foundations of the economics of innovation – theory, measurement and practice. Cheltenham: Edward Elgar Publishing.

Kearnes, D.T. (1986). Quality improvement begins at the top, World, 20 (5), 21.

Kodama, F. (1987). A system approach to science indicators. In Grupp (Ed.), Problems of measuring technological change (pp. 65-87). Köln: Verlag TÜV Rheinland.

Management and Coordination Agency (Ed.). (various years). Report on the survey of research and development, Tokyo.

National Science Board (Ed.). Science (& engineering) indicators, Washington, D.C.

Narin, F., Noma, E. (1985). Is technology becoming science? Scientometrics, 7, 369–381.

Niwa, F., Tomizawa, H. (1995). Integrated indicators: international comparisons of overall strengths in Science and Technology. In: NISTEP (Ed.), Science and Technology Indicators 1994 (pp. 345–365). Tokyo: Science and Technology Agency.

OECD (Ed.). (2003). OECD Science, Technology and Industry Scoreboard. Paris: Organisation for Economic Co-operation and Development.

OST (Ed.). (various years). Science & Technologie – Indicateurs. Paris: Economica.

Patel, P., Pavitt, K. (1995). Patterns of technological activity: their measurement and interpretation. In Stoneman (Ed.), (pp. 15–51).

Pavitt, K. (1988). Uses and abuses of patent statistics. In: A.F.J. van Raan, Handbook of quantitative studies of science and technology (pp. 509–536). Amsterdam: North Holland.

Pohn-Weidinger, S., Gassler, H., Gruber, M., Polt, W., Schibany, A., Leo, H. (2001), Innovationsbericht 2001. Wien: Bundesministerium für Wirtschaft und Arbeit.

Republik Österreich (Ed.). (2003). Österreichischer Forschungs- und Technologiebericht 2003. Wien: Bericht der Bundesregierung an den Nationalrat.

Stoneman, P. (Ed.). (1995). Handbook of the Economics of Innovation and Technological Change. Oxford and Cambridge, Mass.: Blackwell.

Tijssen, R.J.W. (2003). Scoreboards of research excellence. Research Evaluation, 12 (2), 91–103.

Chapter 4

KEEPING THE GATES OF SCIENCE JOURNALS

Gatekeeping Indicators of National Performance in the Sciences

Tibor Braun

Information Science and Scientometrics Research Unit (ISSRU), Institute for Research Organization, Hungarian Academy of Sciences, Budapest, Hungary; and Institute of Inorganic and Analytical Chemistry, L. Eötvös University, Budapest, Hungary
E-mail: braun@helka.iif.hu

Abstract: The chapter outlines the basics of journal gatekeeping and stresses the gatekeeping indicators initiated by us in the 1980s for the evaluation of the research performance of nations in all fields of science. The results have proven that these indicators correlate well with science indicators built on journal paper productivity and citations of nations and represent implicitly a component of quality somewhat associated with citation impact of journal papers. That is why gatekeeping indicators are useful and more simple to build as indicators based on publication productivity and citation impact.

1. INTRODUCTION

The present system of basic research in the sciences and scientific communication depends almost entirely on the primary journal literature. Modern science has developed a particular mechanism of communication which began with the appearance of the first scientific journals in the 17th century and which has remained basically the same ever since. Briefly, this mechanism is based on the selective publication of fragments rather than complete treatises. It is this selective concern with fragments of knowledge, represented primarily by journal articles, that enables science to function effectively and is responsible for its phenomenal growth and pre-eminence (Zuckerman, 1971).

That a paper has been accepted for publication in a well known refereed journal is probably the best immediate indication that it reports worthwhile

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 95-114.

research. This approach is based on the assumption that the primary literature represents the only genuine record of scientific achievement.

For the satisfactory operation of this international mechanism in the sciences the control and screening activity of journal editorial boards, which guarantee the professional standard of science journals, is of paramount importance. It is considered, that the critical mentality and decisions of journal editors have so far protected and will also warrant in the future the social and intellectual integrity of science. The members of the editorial and advisory boards of science journals are rightly considered the gatekeepers of the science journals. These gatekeepers, in controlling the systems of manuscript evaluation and selection, occupy powerful strategic positions in the collective activity of science (Crane, 1967). Taking into account their vital strategic importance in the orchestration of science, it seems interesting to have some quantitative data on the journal gatekeeping process (Zsindely, 1982; Braun, 1983; Zsindely, 1989).

2. METHODS AND RESULTS

Science journals can be considered 'international' if their editorial board included scientists from e. g., five countries at least, irrespective of the title of the journal in question. (The 'international' label in the title of some journals may hide a truly national journal. In contrast, in the editorial board of, e.g., the American Heart Journal there are, in addition to North Americans, scientists from ten, mostly European, countries.)

The classification of journals by fields followed that used previously (Braun, 1985).

The necessary data were obtained by counting and pooling the editors according to country. In so doing, we considered editors, the editor–in–chief, the editor(s), the deputy editor(s) (in–chief), the managing editor, the members of the editorial board and advisory board, excepting only the technical editor(s).

Table 4.1 shows the field, country and geopolitical regional distribution of editors in a sample of 252 journals.

It shows quite clearly that the decision power in science journal gate-keeping is firmly in the hands of scientists from a few (4–5) developed countries. As a group, gatekeepers from the South play a very modest role in the decision power of science journals.

The data can be correlated with the number of science papers published by authors of the respective geopolitical regions and with the number of science journals (Carpenter, 1980). The correlations were investigated within each field. A correlation study among the number of first authors and the number of gatekeepers for each country supplemented those enumerated.

The correlation coefficients and the slopes of the regression lines (calculated by the method of unweighted least squares) were in each case determined from double logarithmic plots.

Table 4.1a. Editorial gatekeeping patterns in a selected set of science journals

Country / Region	*All fields*		*Clinical med*		*Biomed res*		*Chemistry*		*Physics*	
	Rk[a]	%	Rk	%	Rk	%	Rk	%	Rk	%
USA	1	28.8	1	30.0	1	30.2	1	24.6	1	25.7
UK	2	14.6	3	14.0	2	16.2	2	14.0	2	21.4
Rest of W. Europe (WEU)	3	13.6	2	18.2	3	13.6	3	12.3	3	12.8
FR Germany (FRG)	4	10.2	4	9.0	4	13.2	4	11.3	4	10.3
France (F)	5	5.3	7	4.2	5	5.0	5	7.6	5	7.2
East Europe (EEU)	6	4.7	6	4.3	10	2.5	6	6.9	6	5.6
Japan (JAP)	7	3.1	9	2.4	7	2.6	7	4.6	8	2.5
Italy (I)	8	2.9	10	2.4	13	1.5	9	3.1	10	2.0
Canada (CDN)	9	2.8	8	2.9	11	1.6	10	2.8	9	2.5
Soviet Union (SU)	10	2.7	11	1.5	8	2.6	8	3.4	7	2.5
Sweden (S)	11	2.2	5	4.3	9	2.6	13	1.4	11	1.7
Australia (AUST)	12	1.9	12	1.3	12	1.5	11	1.5		
German DR (DDR)	13	1.6	14	0.7	6	3.3				
Israel (IL)	14	1.2	15	0.7	14	1.2	12	1.5		
Latin America (LAM)	15	1.0	13	0.9			15	0.8		
India (IND)	16	0.9	16	0.6			14	0.9		
South Africa (SAF)	17-18	0.4								
Rest of Near East & N Africa (NREA)	17-18	0.4								
Rest of Asia (ASP)	19	0.3								
Other		1.1		2.6		2.6		1.9		5.8
No. of journals		252		45		28		22		49
No. of editors		8,222		1,742		937		615		1,688

[a] Rk: Rank. For Table 4.1b with data for four other fields see next page

Table 4.2 shows the correlation coefficient (r) and the slope (m) of the regression line between the logarithm of the number of gatekeepers and the logarithm of the number of papers or the logarithm of the number of journals published in each field. Correlation parameters between the sum of the gatekeepers in all fields and the number of authors of the given regions were determined as well, again on double logarithmic scales.

Table 4.1b. Editorial gatekeeping patterns in a selected set of science journals

Country / Region	*Biological*		*Earth Space sci*		*Eng. Techn.*		*Mathem*	
	Rk	%	Rk	%	Rk	%	Rk	%
USA	1	21.5	1	21.8	1	35.6	1	32.4
UK	4	11.2	3	9.5	2	15.6	4	6.8
Rest of W. Europe (WEU)	2	15.1	2	17.8	3	9.7	2	13.0
FR Germany (FRG)	3	11.9	4	8.9	4	9.2	5	6.5
France (F)	9	3.9	5	7.4	6	4.0	7	4.0
East Europe (EEU)	7	4.7			5	4.1	6	5.6
Japan (JAP)	10	2.9			8	3.0		
Italy (I)	6	5.4			11	1.9	3	13.0
Canada (CDN)	8	4.6	6	6.9	9	2.4		
Soviet Union (SU)	12	2.1			7	3.4		
Sweden (S)			9	3.4	14	1.3		
Australia (AUST)	11	2.6	7	5.7	10	2.3		
German DR (DDR)	5	5.4	8	3.4				
Israel (IL)					13	1.6		
Latin America (LAM)					12	1.7		
India (IND)								
South Africa (SAF)					16	0.7		
Rest of Near East & N Africa (NREA)					15	1.1		
Rest of Asia (ASP)					17	0.6		
Other		8.7		15.2		1.1		18.7
No. of journals		*25*		*10*		*59*		*14*
No. of editors		*709*		*349*		*1,858*		*324*

Table 4.2. Correlation parameters for the relationship between the number of gatekeepers and the number of science journals, the number of authors, and the number of papers of each field and for all fields combined

Field	*G-P*		*E-J*		*G-A*	
	r	*m*	*r*	*m*	*r*	*m*
Clinical medicine	0.948	1.258	0.757	0.883	—	—
Biomedical Research	0.901	1.135	0.803	0.828	—	—
Biology	0.629	0.651	0.811	0.939	—	—
Chemistry	0.910	1.082	0.851	0.965	—	—
Physics	0.872	0.973	0.839	0.988	—	—
Earth and Space Sciences	0.659	0.697	0.713	0.823	—	—
Engineering and Technology	0.880	0.754	0.839	0.749	—	—
Mathematics	0.750	0.691	0.716	0.794	—	—
Total	0.913	1.079	0.808	0.997	0.899	0.924

G–P = log number of gatekeepers – log number of papers; G–J = log number of gatekeepers – log number of journals; G–A = log number of gatekeepers – log number of authors

A strong correlation is observed between the number of gatekeepers and the number of papers for each scientific field (Table 4.2). The correlation coefficients for the particular fields are above 0.650, and for all fields combined they are above 0.800.

The slopes of the regression lines are between 0.651 and 1.258. Using the number of gatekeepers of all science journals, the regression coefficients are close to unity (1.079 and 0.997, respectively). This suggests direct linear relationships between the number of gatekeepers and papers, and the number of gatekeepers and journals. For the regression among the total number of gatekeepers and the number of authors the slope is again close to unity.

The position of the individual points (corresponding to the single geopolitical regions) relative to the regression line between the number of gatekeepers and the number of journals suggests the following. Countries or geopolitical regions situated above the regression line have more gatekeepers in the editorial boards of international journals than may be expected from the number of science journals published in the country or geopolitical region in question. For the countries below the regression line the situation is opposite.

The country–by–country distribution of the number of gatekeepers and journals in the various fields shows that the US, Canada, Israel, Sweden, Germany, the United Kingdom and France are situated above the regression line, whereas Japan is almost always situated below it.

In most of the science fields the distribution of the number of gatekeepers and papers shows that the points corresponding to Sweden, Germany and Western Europe are above, while those for India and Japan are below the regression line.

Upon analysing the totality of gatekeepers in the whole sample relative to the number of journals one finds more international journal gatekeepers than expected in the case of Israel, Sweden, Canada, France, Germany, the rest of Western Europe, leaving out the United Kingdom and the United States, while less than expected are found for Japan and India. Related to the number of papers, the number of gatekeepers is higher than expected in Sweden, Italy, Eastern and Western Europe, Germany, whereas the number of gatekeepers from Canada and Japan is lower than expected. An essentially identical statement can be made relating the number of gatekeepers of all science journals covered to the number of SCI authors.

Along with the scientific development level of each country one should also take into account, the 'open' or 'closed' nature of scientific life in the country in question. In other words, one should also consider how actively scientists of a country take part in the international scientific life, or to what extent are they isolated (by language or other barriers).

There have earlier been some efforts to answer the above question. Frame and Carpenter (1979) studied the international co-authorship pattern in scientific papers. Previous investigations on international publication practices have pointed out that only 12% of the Israeli researchers publish their results in domestic journals, and also scientists from The Netherlands, Japan and Switzerland publish preferentially in foreign, and in particular in US journals.

In the case of some countries, the geographical distance could result in their scientists figure appearing less frequently among the members of the editorial boards of international journals (Australia, New Zeeland).

The countries whose scientific life is more open are in a better position. Among them Israel, Sweden, Canada, the United Kingdom, and the U.S. belong to this category. Due to their efficient international relationships (e.g., co-authorship with foreign researchers) and to their more successful communication strategy, etc., the scientists of these countries are more 'visible' and so they receive relatively more invitations to participate in the editorial boards of international journals.

Another important question is that of to what extent the status of the gatekeepers of international chemistry journals influences the quality of journals they gatekeep. Additionally, it is also important to clarify whether it is the editor(s)–in–chief alone or the full editorial and advisory board that influences decisively the professional standing of science journals. The editor(s)–in–chief can be identified using Ulrich's Directory (1979).

The professional impact of the gatekeepers can be measured by the number of citations to all of their previously published papers. These figures can be compared only within the same science field, as citation practices and behaviour vary among different disciplines.

As a measure of the weight of international journals, their impact factors can be used, as given in the Journal Citation Reports. The use of this measure, being based on the citation frequency of papers published in the given journals, is consistent with the choice of using the number of citations as a measure of the professional impact of the gatekeepers.

The influence of gatekeepers and editor(s)–in–chief upon the quality of the journals they gatekeep, can be estimated from a correlation between the per capita citations to the gatekeepers and the impact factors of the journals in question.

A next goal is to establish whether the number of citations to the gatekeepers is a valid science indicator. To this end, correlations were set up between the number of citations to gatekeepers of various nationalities, and the number of chemical journals published, or of papers published in the chemical field, in the respective country. Correlations were also sought between the number of gatekeepers from various countries, and their citation frequencies.

The correlation coefficient and the slope of the regression line obtained by the method of unweighted least squares were always determined between the logarithms of the variables in question.

Table 4.3 shows the impact factors of 49 chemistry journals and the mean citation frequency of their gatekeepers. The number of all gatekeepers in the table includes the editor(s)–in–chief as well.

Table 4.3. The impact factor and the mean citation frequency of gatekeepers of international chemistry journals.

Title of journal	*Impact factor*	*Editor(s)–in–chief*			*All editors*		
		No.	*citation freqn.*	*per capita cit. freqn.*	*No.*	*citation frequn.*	*per capita cit. freqn.*
Acta Crystallographica	1.133	1	989	989	15	4,070	271
Advances in Colloid and Interface Science	1.368	2	585	293	21	4,540	217
Analytical Letters	0.884	1	952	952	61	14,519	238
Analusis	0.774	1	262	262	49	5,907	121
The Analyst	1.702	—	—	—	42	8,664	206
Analytica Chimica Acta	1.488	1	88	88	39	7,707	198
Carbohydrate Research	1.431	—	—	—	53	6,638	125
Chromatographia	1.394	—	—	—	33	8,978	272
Electrochimica Acta	1.048	1	18	18	18	2,668	148
European Polymer Journal	1.044	1	125	125	22	4,138	188
Fluoride	0.705	1	123	123	27	1,603	59
Inorganica Chimica Acta	2.859	1	44	44	78	42,086	540
Inorganic and Nuclear Chemistry Letters	1.141	1	412	412	25	14,029	561
International Journal of Chemical Kinetics	1.959	1	2,609	2,609	19	9,005	474
International Journal of Polymeric Materials	0.720	1	15	15	35	4,270	122
International Journal for Radiation Physics and Chemistry	—	—	—	—	18	5,312	295
Journal of Applied Crystallography	0.861	1	930	930	6	840	140
Journal of Chemical Technology and and Biotechnology	—	—	—	—	27	1,957	72
Journal of Chromatography	1.846	1	265	265	45	11,278	251
Journal of Computational Chemistry	—	—	—	—	18	24,056	1,336
Journal of Inorganic and Nuclear Chemistry	1.017	1	412	412	72	28,223	392

Title of journal	*Impact factor*	*Editor(s)–in–chief* No.	*citation freqn.*	*per capita cit. freqn.*	*All editors* No.	*citation frequn.*	*per capita cit. freqn.*
Journal of Macromolecular Science, Chemistry	0.440	1	252	252	43	10,274	247
Journal of Organometallic Chemistry	2.331	—	—	—	7	9,888	1,413
Journal of Radioanalytical Chemistry	0.890	2	129	65	47	4,470	95
Journal of Raman Spectroscopy	0.900	1	341	341	42	12,801	312
Journal of Thermal Analysis	0.506	2	48	24	32	3,601	113
Kristall und Technik	—	1	2	2	30	3,491	116
Die makromolekulare Chemie	1.140	1	274	274	64	16,486	257
Mikrochimica Acta	0.779	1	5	5	41	8,825	215
Molecular Crystals and Liquid Crystals	1.016	—	—	—	30	12,774	426
Monatshefte für Chemie	0.831	—	—	—	38	13,584	357
Organic Magnetic Resonance	1.379	1	15	15	38	16,538	435
Organic Mass Spectrometry	1.253	1	354	354	36	14,824	411
Pure and Applied Chemistry	1.433	—	—	—	7	992	142
Radiochimica Acta	0.573	1	—	—	13	1,868	144
Radiochemical and Radioanalytical Letters	0.515	2	201	101	72	6,445	90
Spectrochimica Acta, Part A	1.023	1	49	49	33	5,540	168
Spectrochimica Acta, Part B	1.621	1	37	37	32	1,5490	470
Starch — Stärke	0.646	1	28	28	10	2,980	298
Synthesis and Reactivity in Inorganic and Metalorganic Chemistry	0.905	1	344	344	41	22,624	552
Synthesis	1.758	—	—	—	24	27,026	1,126
Synthetic Communications	1.178	1	936	936	29	17,424	602
Talanta	0.907	1	195	195	14	3,900	279
Tetrahedron	1.745	1	557	557	70	5,9728	853
Tetrahedron Letters	2.114	1	557	557	64	5,9540	930
Texture of Crystalline Solids	—	1	234	234	23	5,301	230
Theoretica Chimica Acta	1.816	1	310	310	21	12,194	581
Thermochimica Acta	0.675	1	519	519	27	4,190	155

A strong positive correlation can be found between the mean citation frequency of the full editorial board and the impact factor of the journal. The value of the correlation coefficient between the logarithms of the variables is 0.627 which corresponds to a 99% significance level. When only the citation frequency of the editor(s)–in–chief is used a much lower correlation coefficient (0.218) is obtained.

The discrepancy between these two correlations is quite obvious and suggests that the editor(s)–in–chief perform their gatekeeping function in close cooperation with the members of the editorial and advisory boards of their journals rather than alone. The collective effect of all gatekeepers upon the impact factor of the journals appears to be stronger.

There is a correlation between the number of citations to the gatekeepers of various nationalities on the one hand, and the number of these gatekeepers on the other. Similarly, correlations between the number of citations to the gatekeepers and the number of chemistry journal papers or chemistry journals published in the given country can be seen in Table 4.4.

The correlation coefficients between the logarithms of the variables are high (between 0.807 and 0.940). The slope of the regression lines is always greater than unity, which indicates a slightly non-linear relationship among the original variables. This is yet another proof of the 'citation breed citations' phenomenon. Those who already enjoy a high citation frequency are likely to be cited more frequently in the future.

Table 4.4. Correlation parameters between the citation frequency of gatekeepers of chemical journals on the one hand, and the number of gatekeepers, papers, and journals published, resp., on the other

	Cit-G	*Cit-P*	*Cit-J*
r	0.940	0.876	0.807
m	1.179	1.307	1.146

Cit-G: log of No. of citations - log of No. of gatekeepers; Cit-P: log of No. of citations - log of No. of papers; Cit-J: log of No. of citations - log on No. of journals.

As expected, the best fit is obtained between the number of gatekeepers and their citation frequency (r = 0.941). The dispersion of the individual countries around this regression line differs, however, from that observed in the correlations between the number of gatekeepers and the number of papers and journals. Thus — for the latter correlation — Japan and India are found below the regression line, whereas these countries are above it in the former case. France, some non-specified East- and West-European countries are also found below the regression line, together with countries involving geopolitical regions of developing countries. Compared to earlier findings about the number of gatekeepers, this supports the view that albeit the share of these countries in the gatekeepers board of chemical journals is less than expected on the basis of other science indicators, the citation frequency of their scientists co-opted in the editorial boards of international chemistry journals is, on the average, greater than for scientists of other countries. To put it differently, in order to qualify for an invitation to the editorial advisory board of an international chemistry journal a scientist from Japan or India has to be more 'visible', reflected in his citation frequency than one from France or those of East or West European countries. When the number of citations to the gatekeepers is plotted against the number of papers, or the number of journals published in each country (on double logarithmic scale again) the patterns are similar to those obtained with the number of gatekeepers.

The impact factors of chemistry journals differ over about the same range as the citation rates of their gatekeepers. Do the scientific quality and distinction of the gatekeepers have a repercussion upon their gatekeeping activities in different subfields of chemistry?

An answer to this question can be given by comparing the impact factors of the journals with the citation rates of their gatekeepers. Tables 4.5 to 4.7 contain data for organic, inorganic and analytical chemistry journals, whereas the average impact factors are almost the same for the organic and inorganic chemistry journals and that for the analytical journals is only about 25% lower (Table 4.8). These differences in impact factors are not significant. Between the specific citation rates of the gatekeepers and the impact factor of their journals there is a significant correlation ($r = 0.6$). The slope of the regression line is 0.4, which means that the prestige of journals is only slightly raised by increasing the prestige of the gatekeepers.

Table 4.5. Impact factors of organic chemistry journals and citation data for their gatekeepers

Journal	*Impact factor*	*Gatekeepers*		
		No.	*Total citations*	*Citations per capita*
Carbohydrate Research	1.431	53	6,638	125
Journal of Organometallic Chemistry	2.331	7	9,888	1,413
Monatshefte für Chemie	0.831	38	13,584	357
Organic Magnetic Resonance	1.379	39	16,553	424
Organic Mass Spectrometry	1.253	37	15,178	410
Synthesis	1.758	24	27,026	1,126
Synthetic Communications	1.178	30	18,360	612
Tetrahedron	1.745	71	60,285	849
Tetrahedron Letters	2.114	65	60,097	925

Table 4.6. Impact factors of inorganic chemistry journals and citation data for their gatekeepers

Journal	*Impact factor*	*Gatekeepers* *No.*	*Total citations*	*Citations per capita*
Inorganica Chemica Acta	2.859	79	42,130	533
Inorganic and Nuclear Chemistry Letters	1.141	26	14,441	555
Journal of Inorganic and Nuclear Chemistry	1.017	73	28,635	392
Zeitschrift für anorganische und allgemeine Chemie	1.333	38	15,220	400

Table 4.7. Impact factors of analytical chemistry journals and citation data for their gatekeepers

Journal	*Impact factor*	*Gatekeepers* *No.*	*Total citations*	*Citations per capita*
Analytical Chemistry	2.803	17	3,193	188
Analytical Letters Parts A and B	0.884	62	15,471	250
Analusis	0.774	50	6,169	123
The Analyst	1.702	42	8,664	206
Analytica Chimica Acta	1.488	40	7,795	195
Chromatographia	1.394	33	8,978	272
Journal of Chromatography	1.846	46	11,543	251
Journal of Radioanalytical Chemistry	0.890	49	4,535	93
Journal of Thermal Analysis	0.506	34	3,625	107
Mikrochimica Acta	0.779	42	8,830	210
Radiochemical and Radioanalytical Letters	0.515	74	6,546	88
Spectrochimica Acta, Part A	1.023	34	5,589	164
Spectrochimica Acta, Part B	1.621	33	15,527	471
Talanta	0.907	51	10,831	212

Table 4.8. Comparison of organic, inorganic and analytical chemistry journals

Characteristics	*Organic chemistry*	*Inorganic chemistry*	*Analytical chemistry*
Average impact factor	1.56 ± 0.47	1.59 ± 0.85	1.22 ± 0.62
Average number of gatekeepers per journal	40 ± 20	54 ± 26	43 ± 14
Average citations per gatekeeper	693 ± 415	470 ± 85	202 ± 97

As mentioned, participation in gatekeeping for some scientific journals represents a form of reward for the person involved. Participation in the editorial board of many journals is naturally an accumulated reward, and in such cases, no doubt the ‘Matthew effect’ is at work. It has been shown that scientists who are already known, i.e., more 'visible', are given more reward than others who may have similar scientific achievements but are less 'visible' and/or less widely known.

Among e.g., the 608 gatekeepers of 14 analytical chemistry journals considered, 61 are members of two editorial boards and 19 participate in three or more.

The citation rate of gatekeepers of analytical chemistry journals can be described well by a logarithmic normal distribution curve (Figure. 4.1). The median corresponds to M = 100 citations per 5 yr; in other words, 50% of the gatekeepers receive over 20 citations per year, whereas 68% of them get between 3 and 100 yearly citations ($M \pm \sigma$).

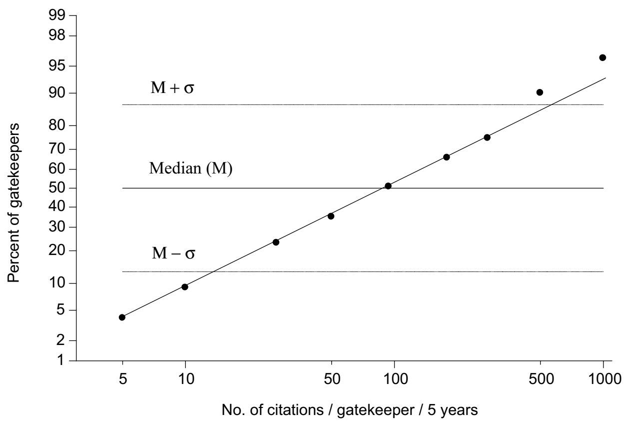


Figure 4.1. Distribution of the number of citations for gatekeepers of journals in analytical chemistry

The recent research of the authors group in using gatekeeping indicators in analytical chemistry and comparing data from the eighties with those at the end of the millennium has proved once again the usefulness of the gatekeeping indicators (Braun, 2004). Investigation are on their way in extending gatekeeping indicators to the evaluation of the scientific weight of nations in 12 science fields during the last decade of the millennium (Braun, 2004).

An additional attempt gives an even more inquisitive response about the professional status and influence of the editor(s)–in–chief of 769 medical journals. Answers were sought to the question of whether the editors–in–chief as authors have larger influence and/or authority than an average scientist in the respective subject field.

Two main indicators were built (citation window: 5 years of publication and citations for the same period):

– An Indicator of Editor Expertise (IEE): the ratio of the editor–in–chief's mean citation rate per cited paper to that of his or her journal;
– An Indicator of Editor Authority (IEA): the ratio of the editor–in–chief's percentage of in-journal citations to that of his or her journal.

Table 4.9. Mean citation rate per cited paper and the Indicator of Editor Expertise (IEE) in medical subfields

Subfield	*Number of editors*	*Mean citation rate per cited paper*		*IEE*
		Journals	*Editors*	
Allergy	8	3.46	3.46	1.00
Andrology	5	3.02	2.57	0.85
Anesthesiology	8	3.53	3.33	0.94
Cancer	48	4.72	3.39	0.72
Cardiovascular system	46	5.30	2.72	0.51
Dentistry and odontology	21	2.86	2.00	0.70
Dermatology and venereal diseases	18	3.17	2.60	0.82
Endrocrinology and metabolism	46	6.09	4.27	0.70
Gastroenterology	19	4.15	2.11	0.51
General and internal medicine	65	2.89	2.26	0.78
Geriatrics and gerontology	10	2.61	2.66	1.02
Hematology	32	5.81	3.10	0.53
Immunology	66	6.02	3.13	0.52
Neurosciences	106	5.47	3.13	0.57
Obstetrics and gynecology	27	3.56	2.10	0.59
Ophthalmology	15	3.53	2.20	0.62
Orthopedics	9	2.45	2.12	0.87
Otorhinolaryngology	9	2.61	1.87	0.71
Pathology	39	4.61	2.96	0.64
Pediatrics	31	2.92	2.59	0.89
Psychiatry	36	3.72	2.31	0.62
Radiology and nuclear medicine	40	3.64	1.90	0.52
Research and experimental medicine	28	3.77	3.02	0.80
Respiratory system	19	4.22	2.62	0.62
Rheumatology	13	2.68	2.78	1.04
Surgery	58	3.02	1.92	0.63
Tropical medicine	12	2.71	2.57	0.95
Urology and nephrology	21	2.81	2.78	0.99

Both indicators have a value of 1.00 if there are no specific differences between the citation record of the editor–in–chief and an average author. All editors–in–chief having at least one cited paper in the period in question (709 persons, 855 editorial chairs) were included in the determination of the Indicator of Editor Expertise (IEE); all editors–in–chief having at least one cited paper in their own journals in the period in question (353 persons, 435 editorial chairs) were included in the determination of the Indicator of Editor Authority (IEA).

Table 4.10. Percentage of in-journal citations and the Indicator of Editor Authority (IEA) in medical subfields

Subfield	*Number of editors*	*Percentage of in-journal citations*		*IEA*
		Journals	*Editors*	
Allergy	5	20	29	1.44
Andrology	5	13	25	1.92
Anesthesiology	6	25	69	2.79
Cancer	21	14	14	1.00
Cardiovascular system	25	13	26	1.94
Dentistry and odontology	12	29	38	1.30
Dermatology and venereal diseases	14	25	32	1.27
Endrocrinology and metabolism	20	14	12	0.84
Gastroenterology	9	10	26	2.56
General and internal medicine	35	21	50	2.33
Geriatrics and gerontology	5	20	52	2.52
Hematology	21	12	21	1.72
Immunology	30	15	23	1.58
Neurosciences	51	16	22	1.38
Obstetrics and gynecology	8	12	16	1.38
Ophthalmology	10	21	18	0.86
Orthopedics	4	11	7	0.63
Otorhinolaryngology	7	19	49	2.56
Pathology	18	13	26	1.94
Pediatrics	19	20	35	1.70
Psychiatry	17	17	30	1.77
Radiology and nuclear medicine	15	21	39	1.87
Research and experimental medicine	16	17	35	2.07
Respiratory system	10	15	13	0.92
Rheumatology	7	17	5	0.27
Surgery	29	18	23	1.29
Tropical medicine	6	23	39	1.67
Urology and nephrology	10	17	29	1.74

The overall average IEE value was 0.59, the overall average IEA value was 1.64. (The subset of editors considered in evaluating IEA had an average IEE of 0.61, i.e., no significant difference from the total set has been found.) The values of both indicators are presented at a subfield aggregate level in Tables 4.9 and 4.10, respectively.

Subfield differences among IEE and IEA indexes, although interesting to consider, are, in general, not always statistically significant.

3. CONCLUSIONS

The abovementioned results can be summarized as follows:

1. In the case of science journals a correlation has been shown to exist between the number of gatekeepers of a given nationality and the number of papers published in these journals by scientists in the country concerned.
2. For science journals a positive correlation also exists between the number of gatekeepers and their citation rates.
3. The relationship between the number of gatekeepers (*n*) and their publication per capita and their citedness rate, *N* is $n \sim aNm$, where m shows values between 0.6 and 0.8. In other words, for the journals mentioned so far the effort needed for a country to increase the number of gatekeepers by one, say from 50 to 51, or from 100 to 101, would be twice and thrice, respectively, as large as that necessary to effect an increase from 10 to 11.
4. It seems that 75% of the positions of power influencing the publication of new results in almost all areas of science are concentrated in the hands of scientists from no more than ten countries of the world.
5. In trying to answer the question of whether the editors–in–chief of medical journals are experts, authorities, both or neither, the main inference to be drawn from the data presented in Tables 4.9 and 4.10 is obvious. In all but 3 of the 28 subfields of medicine, the editors–in–chief are, on average, less cited than the average author in their own journals; and in all but 6 subfields, the average percentage of in-journal citations is significantly higher for the papers of the editor–in–chief than for those of an average author. The answer of the abovementioned question is thus clear: the editors–in–chief are not necessarily experts (in the sense of higher–than–average citation rate) but, as a rule, authorities — at least in their own specialties.
6. The question now arises of if not their research eminence then what else might be the source of the authority of these scientists? An obvious explanation would personal influence, ability to make quick, intuitive decisions, (and so on) may prevent him or her from being a universally acknowledged, highly cited researcher. Of course, the most fortunate cases are those in which the two sets of qualities coincide, but this is the exception rather than the rule.
7. The results stress that the indicators based on journal gate keeping data can be used as usefully as indicators based on journal papers and citations in evaluating national scientific performance.

8. Data based on our pioneering studies and performed also by other authors (Bakker, 1985; Le Minov, 1989; Rigter, 1986; Sievert, 1989; Nisonger, 2002) have shown that in addition to the evaluation of national performance, gate keeping indicators can be successfully used also for evaluations at institutional level. They specifically measure a form of scientific power, expected to be strongly correlated with citation indicators or also excellence measures of various kinds.
9. It is an open question of whether the co-existence of the traditional (paper) and new forms (electronic) of communication will affect gatekeeping in general and implicitly gatekeeping indicators. As this author believes that these forms are not working against each other but they are, and will, coexist synergetically, the gatekeeping will remain as important as it is now in the filtering of new information and in assuring of criterion of new knowledge.
10. All these results suggest that the building of an up to date, comprehensive computerized database of science journal gatekeepers and the continuous maintenance of that database would be a worthwhile addition to the already available series of scientometric tools.

REFERENCES

Bakker, P., Rigter, H. (1985). Editors of medical journals — who and from where. Scientometrics, 7, 11–22.

Braun, T., Bujdosó, E. (1983). Gatekeeping patterns in the publication of analytical-chemistry research. Talanta, 30, 161–167.

Braun, T., Dióspatonyi, I. (2004). The main players in the international gatekeeping orchestration of analytical chemistry journals. Gatekeeping indicators. J. Am. Soc. Inf. Sci. Technol., in print.

Braun, T., Dióspatonyi, I. (2004). The sciencific weight of nations measured via gatekeeping indicators, unpublished data.

Braun, T., Glänzel, W., Schubert, A. (1985). Scientometric indicators. A 32 Country comparison of publication productivity and citation impact. World Scientific Publ. Co., Ltd. Singapore.

Carpenter, M.P., Narin, F. (1980). Subject composition of the Worlds scientific journals. Scientometrics, 2, 53–63.

Crane, D. (1967). The gatekeepers of science: some factors affecting the selection of articles for scientific journals. The American Sociologist, 195–201.

Frame, J.D., Carpenter, M.P. (1979). International research collaboration. Soc. Stud. Sci., 9, 481–497.

Le Minov, S., Dostatni, P. (1989). INSERM Report.

Nalimov, V.V., Mulchenko, Z.M. (1969). Naukometriya. Nauka. Moscow.

Rigter, H. (1986). Evaluation of performance of health research in the Netherlands. Research Policy, 15, 33–48.

Sievert, M.E., Haughawout, M. (1989). An editor's influence on citation patterns — a case-study of elementary-school journal. J. Am. Soc. Inf. Sci., 40, 334–341.
Ulrich's International Periodicals Directory, 18. ed., 1979–1980. New York, London: R.R. Bowker.
Nisonger, T.E. (2002). The relationship between international editorial board composition and citation measures in political science, business, and genetics journals. Scientometrics, 54, 257–268.
Zuckerman, H., Merton, F.K. (1971). Patterns of evaluation in science — institutionalisation, structure and functions of referee system. Minerva, 9, 66–100.
Zsindely, S., Schubert, A. (1989). Editors–in–chief of medical journals — are they experts, authorities, both, or neither. Commun. Res., 16, 695–700.
Zsindely, S., Schubert, A., Braun, T. (1982). Editorial gatekeeping patterns in international science journals — a new science indicator. Scientometrics, 4, 57–68.
Zsindely, S., Schubert, A., Braun, T. (1982). Citation patterns of editorial gatekeepers in international chemistry journals. Scientometrics, 4, 69–76.

APPENDIX

Table 4.A1. Journals included into the analyses per discipline

Journal Title	*Journal Title*
Clinical Medicine	***Biomedical Research***
Acta Allergologica	Archives of Microbiology
Acta Endrocrinologica	Archives of Virology
Acta Oto-Laryngologica	Behavior Genetics
American Heart Journal	Biochemical Genetics
Atherosclerosis	Biochemical Systematics and Ecology
Biochemical Pharmacology	Biologischer Zentralblatt
Biology of the Neonate	Biometrical Journal
Clinical Orthopaedics and Related Research	Biosystems
Digestion	Biotelemetry and Patient Monitoring
Environmental Research	Bulletin of Mathematical Biology
European Journal of Cancer	Chemistry and Physics of Lipids
European Journal of Clinical Pharmacology	Chromosoma
Experimental Gerontology	European Journal of Applied Physiology and Occupational Physiology
Geriatrics	European Journal of Biochemistry
Gerontology	European Journal of Physiology
Gynecologic and Obstetric Investigation	Experimental Cell Research
Haematologia	International Journal for Vitamin and Nutrition Research
Immunochemistry	Journal of General Microbiology
International Archives of Allergy and Applied Immunology	Journal of Molecular Biology
International Archives of Occupational and Environmental Health	Journal of Theoretical Biology
International Journal of Applied Radiation and Isotopes	Medical and Biological Engineering and Computing
International Journal of Cancer	Molecular and Cellular Biochemistry
International Journal of Environmental Studies	Molecular and General Genetics

Journal Title	*Journal Title*
International Journal of Radiation Biology and Related Studies in Physics, Chemistry and Medicine	Photochemistry and Photobiology
International Urology and Nephrology	Photoplasma
Journal of Bone and Joint Surgery American	Roux's Archives of Developmental Biology
Journal of Bone and Joint Surgery British	Theoretical and Applied Genetics
Journal of Clinical Pathology	Theoretical Population Biology
Journal für Hirnforschung	
Journal of Investigative Dermatology	***Chemistry***
Journal for Oto-Rhino-Laryngology and its Borderlands	
Lung	Acta Crystallographica
Nuclear Medicine	Advances in Colloid and Interface Science
Oncology	Analusis
Ophthalmologica	Analyst
Psychiatria Clinica	Analytica Chimica Acta
Psychopharmacology	Analytical Letters
Respiration	Carbohydrate Research
Respiration Physiology	Chromatographia
Thrombosis and Haemostasis	Electrochimica Acta
Toxicon	European Polymer Journal
Virchows Archiv, Abt. A	Fluoride
Virchows Archiv, Abt. B	Inorganica Chimica Acta
Vox Sanguinis	Inorganic and Nuclear Chemistry Letters
Water, Air and Soil Pollution	International Journal of Chemical Kinetics
	International Journal of Polymeric Materials
Biology	International Journal of Radiation Physics and Chemistry
	Journal of Applied Crystallography
Agricultural Meteorology	Journal of Chemical Technology and Biotechnology
Agrochimica	Journal of Chromatography
Animal Feed Science and Technology	Journal of Computational Chemistry
Aquaculture	Journal of Inorganic and Nuclear Chemistry
Beiträge zur Entomologie	Journal of Macromolecular Science Chemistry
Biological Cybernetics	Journal of Organometallic Chemistry
Cryobiology	Journal of Radioanalytical Chemistry
Ecological Modelling	Journal of Raman Spectroscopy
Environmental and Experimental Botany	Journal of Thermal Analysis
European Journal of Forest Pathology	Kristall und Technik
Geoderma	Makromolekulare Chemie
Journal of Fish Biology	Microchimica Acta
Journal of Mathematical Biology	Molecular Crystals and Liquid Crystals
Landwirtschaftliches Zentralblatt I. Landtechnik	Monatshefte für Chemie
Milchwissenschaft	Organic Magnetic Resonance
Physiological Plant Pathology	Organic Mass Spectrometry
Phytochemistry	Pure and Applied Chemistry
Radiation and Environmental Biophysics	Radiochimica Acta
Scientia Horticulturae	Radiochemical and Radioanalytical Letters
Zeitschrift für Botanische Taxonomie und Geobotanik	Spectrochimica Acta Part A
Zeitschrift für Pflanzenphysiologie	Spectrochimica Acta Part B
Zeitschrift für Pflanzenzüchtung	Starch
	Synthesis
Physics	Synthesis and Reactivity in Inorganic and MetalOrganic Chemistry
	Synthetic Communications
Acta Physica Austriaca	Talanta

Journal Title
Acta Physica Polonica A
Acustica
Advances in Molecular Relaxation and Interaction Processes
Advances in Physics
Chemical Physics Letters
Cryogenics
Foundations of Physics
Journal of Computational Physics
Journal of Physics C
Journal of Physics and Chemistry of Solids
Journal of Sound and Vibration
Metrologia
Molecular Physics
Nuclear Physics A
Optica Acta
Physica Status Solidi B
Physics and Chemistry of Liquids
Physics Letters C
Revue Internationale des Hautes Températures et des Réfractaires
Semiconductors and Insulators
Surface Science
Thin Solid Films
Ultrasonics
Vacuum
Zeitschrift für Physik B
Engineering and Technology
Acta Informatica
Acta Metallurgica
Aerosol Report
Angewandte Informatik
Annals of Nuclear Energy
Astronautica Acta
Atomic Energy Review
Atomkernenergie
Automatica
Brauwissenschaft
Cement and Concrete Research
Chemical Engineering Journal
Chemical Engineering Science
Combustion Science and Technology
Composites
Computer Aided Design
Computer Graphics and Image Processing
Computer Methods in Applied Mechanics and Engineering
Computers and Electrical Engineering
Computers and Operations Research
Computers and Structures
Corrosion Science

Journal Title
Tetrahedron
Tetrahedron Letters
Texture of Crystalline Solids
Theoretica Chimica Acta
Thermochimica Acta
Zeitschrift für Anorganische und Allgemeine Chemie
Earth and Space Science
Astronomische Nachrichten
Chemical Geology
Geochimica et Cosmochimica Acta
Earth and Planetary Science Letters
Earth Science Reviews
International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts
Mineralium Deposita
Rock Mechanics
Tectonophysics
Tschermaks Mineralogische und Petrographische Mitteilungen
Mathematics
Advances in Mathematics
Algebra Universalis
Archiv für Mathematische Logik und Grundlagenforschung
Internationale Mathematische Nachrichten
Journal of Algebra
Journal of Applied Probability
Journal of Mathematical Analysis and Applications
Logique et Analyse
Mathematics of Computation
Mathematics and Computers in Simulation
Mathematica Scandinavica
Optimalization
Rendiconti di Mathematica
Zeitschrift für Wahrscheinlichkeits-theorie und Verwandte Gebiete

Journal Title	*Journal Title*
Cybernetica	
Desalination	
Earthquake Engineering and Structural Dynamics	
Electrocomponent Scinece and Technology	
Energy Sources	
Engineering Fracture Mechanics	
Fuel	
Information and Control	
Information Sciences	
International Forum on Information and Documentation	
International Journal of Circuit Theory and Applications	
International Journal of Computer and Information Sciences	
International Journal of Control	
International Journal of Electrical Engineering Education	
International Journal of Electronics	
International Journal of Engineering Science	
International Journal of Fracture	
International Journal of Heat and Mass Transfer	
International Journal of Machine Tool Design and Research	
International Journal of ManMachine Studies	
International Journal of Mechanical Sciences	
International Journal of Powder Metallurgy and Powder Technology	
Journal of Computer and System Sciences	
Journal of Hazardous Materials	
Journal of Hydraulic Research	
Journal of Hydrology	
Matériaux et Construction	
Metallography	
Non-Destructive Testing	
Nuclear Engineering and Design	
Powder Technology	
Pulp and Paper International	
Solid State Electronics	
Transportation	
Transportation Researches	
Wärme, Gas-International	
Wear	

Chapter 5

S&T INDICATORS FOR POLICY MAKING IN A CHANGING SCIENCE–SOCIETY RELATIONSHIP

Rémi Barré

Conservatoire National des Arts-et-Métiers (CNAM), Paris, France
E-mail: remi.barre@cnam.fr

Abstract: This paper suggests that an understanding of the changing science–technology–innovation régime can provide a new role for necessarily imperfect S&T indicators. It is argued that it is precisely to the extent which they are questionable that S&T indicators can fulfil their role as mediator and S&T arena decision making instruments. Such a role of S&T indicators is highlighted by initiatives of benchmarking for public policy design and improvement, as shown by recent reports from the European commission and OECD. Finally, it is argued that it has implications for the work and responsibilities of the S&T indicators specialists.

1. INTRODUCTION

The science–society relationship is changing, the innovation related policy making having to deal with ever more complex system of actors, decision makers and stakeholders. Hence, despite progresses, we must admit the very real limitations of S&T indicators, owing to insufficient data as well as inadequate understanding of the functioning of the innovation system. We are more and more aware of the limitation of our knowledge regarding the linkages between research, the economy and society.

Are these facts shattering the very relevance, thus the legitimacy and the future of S&T indicator?

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 115-131.

The aim of this paper is to suggest that an understanding of the changing science–technology–innovation régime can provide a new role for such imperfect S&T indicators in the context of S&T policy making.

2. THE TRANSFORMATION OF THE INNOVATION SYSTEMS (IS) AND ITS MEANING FOR S&T INDICATORS

2.1 The New Science–Society Relationship and its Impact on S&T Indicators

a) The changing status of scientific knowledge

Innovation has become a central feature of the developed industrial societies; indeed, permanent innovation is the norm for a firm to be competitive, and, in fact, for any institution to survive. The new contract between science and society is based on the recognition of both the pervasiveness of innovation and the extraordinary efficiency of science to generate novelty at high speed, which, in turn, generates, at equally high speed, complexity and risk.

The very success of science has put it at the centre of human activity and destiny, which transforms it into a social object. This implies that it cannot anymore hold as an autonomous reality and activity, defined by its own social and professional norms and codes. In this sense science cannot any longer claim to be the unique linkage to universality and rationality (Nowotny et al., 2001; Latour, 1999).

The knowledge claimed as 'scientific' by the scientists no longer has the monopoly of truth in a socio-political context. The scientific expert is placed in a situation where he/she has to argue, to present hypotheses and assumptions not only to his peers, but also in public. The validation of knowledge is realised in a more diversified professional context.

The reason for this is that scientific knowledge, as created by researchers in their laboratory, is very seldom applicable directly in the 'real world' of applications. Put in the wider socio-economic, biological and ecological systems, there are chances that things will not work as expected, that unintended consequences will arise, since the 'all things being equal' used for working in the laboratory, will not apply. In other words, the criteria for scientific truth have not changed. What has changed is that academically sound science is not accepted directly as providing undisputable guidance in the 'real world': it has first to be transformed into 'socially robust' knowledge through a broader validation process (Nowotny et al., 2001).

b) Implications for S&T indicators

The implication for S&T indicators is direct: recent progress in the understanding of innovation systems notably through the development of evolutionary approaches, show that situations are always context specific and path dependent. In other words, there is no unequivocal interpretation to be given to an indicator. Furthermore, the complexity of Innovation systems makes it so difficult to understand the nature of the causal links that no adequate models of such links between these entities exist.

Therefore, even though based on methods having an explicit scientific base, the validity of the indicators is bound to be questioned, and rightly so. No matter how 'scientific' the S&T indicators can be, they will be challenged and become part of the debate rather than pretending to an external anchor to safeguard the so called rationality of the debate.

In the same way as any scientific knowledge to be applied in society, S&T indicators will have to be 'socially robust', that is, to be understood and appropriated by the stakeholders. They will have to pass the test of relevance in a variety of contexts and to a variety of stakeholders.

S&T indicators being imperfect and questionable is not something to be ashamed of, but the norm of any scientific knowledge. To pretend to have the 'true' indicators would be not only a lie, but would also be counter-productive: in the new science–society relationship, such a claim is not receivable any more.

2.2 The Rise of the Agora Model for Innovation and S&T Policy

A general trend of the developed industrial societies consists in the weakening of the separation of the fundamental categories which used to distinguish, for example, the market and the state, private and public sectors, science and values, producers and users of knowledge. In practice, within the 'innovation system' (IS), the actors of the industrial innovation sub-system, the academic sub-system and the socio-political sub-system tend to mix together (Caracostas and Muldur, 1997; Etzkowitz and Leydesdorff, 2000).

In other words, innovation processes as well as research and innovation policy decision making processes, tend to involve increasingly the variety of the components of the innovation system, namely academia, industry, and the citizens. Hence the broadening and complexification of the decision making processes regarding innovation and S&T policy decisions, in terms of players involved, aspects considered, variety of institutional settings and organisational levels to be included.

Such a situation, in which the social actors play an active role in the innovation as well as in S&T policy processes, defines what we call the *Agora model* (Barré, 2001a). It embodies a new science–society relationship in which the quality and efficiency of multi-actors' interaction is crucial for the overall performance of the IS. The Agora model has to do with the *Arena* approach developed by Kuhlmann (Kuhlmann and Edler, 2003a; Kuhlmann, 2003b), although the former concept tends to focus on the debates amongst actors and the latter on the interplay among the various political decision–making levels (regional, national, European).

It is the objective of the *Agora processes* to allow for interactions between industry, academia, government and social actors, that develop in an efficient, deep, and coherent way. They aim at helping the actors to build the hybrid networks through which they will jointly produce knowledge, ensuring also that the knowledge produced is relevant to each sub-system. The best known Agora processes are foresight, technology assessment, socio-economic impact evaluation, and benchmarking, which assume the 'strategic & distributed intelligence' function in the emerging Innovation system (Kulhmann et al., 1999).

Those Agora processes are based on the basic principles (Smits, 2001) of participation (to take care of the diversity of the perspectives of the actors), of mediation and alignment (to facilitate mutual and reflexive learning about the perspective of the different contesting actors), of objectivisation (to support the formulation of diverging perceptions by offering appropriate analysis and information processing mechanisms).

Here comes the implication for S&T indicators: they are central to the implementation of the Agora processes, since there are instruments of objectivisation, allowing as well for participation and mediation.

2.3 S&T Indicators as Agora Processes Instruments

S&T indicators are based on the application of scientific knowledge to the quantitative description of S&T activities and implications. In this sense they are instruments of objectivisation.

But, as we have seen, like any knowledge based on scientific approach, like any form of expertise, S&T indicators do not stand alone any longer, they cannot tell any 'truth' as such. They are questionable and constitute debatable elements related to the central element of the Agora processes, i.e., S&T. In this sense they become an instrument of mediation and participation.

It is precisely to the extent that they are questionable that S&T indicators can fulfil their role as mediator and S&T arena decision making instrument. The indicators become one of the instruments by which the actors shape their debate through the expression of their diverging views about their relevance, significance, underlying assumptions, and interpretation. It is the very process of such criticism which enables actors to address the deeper question related to the decisions at stake. It is the exchange of arguments together with criticism which leads to the revealing of substantive points which are important to the actors, as a mix of facts and values.

This is indeed a crucial role which indicators play here, since numbers have a unique capability of being immediately understood: indicators are the common language amongst actors which allow for the interactions called for by the Agora model. They are highly ‘inter-operable’.

The conclusion of this first part is that the new science–technology–innovation policy régime in the making transforms the S&T indicators into debatable pieces which, far from weakening them, provide them, on the contrary, with a new perspective and *raison d’être*: S&T indicators as elements of the Agora processes (Barré, 2001a).

But how is this role of S&T indicators in the Agora model actually reflected in actual procedures and decision making processes?

3. S&T INDICATORS AS POLICY MAKING INSTRUMENTS IN THE NEW CONTEXT: THE CASE OF BENCHMARKING

The role of S&T indicators in S&T decision making processes has been recently highlighted by initiatives in benchmarking for public policy design and improvement. Benchmarking is a process for systematically comparing performance against ‘the best in the world’ to gather information which helps to take steps towards the improvement of one’s own activity. The concepts behind benchmarking are those of improvement of a situation through identification of best practices through performance assessment and comparisons based on indicators.

In such exercises, obviously, indicators play a central role. In this sense they are objectivisation, participation, and mediation instruments in the context of the emerging innovation system perspective, which we called the Agora model.

3.1 The Open Method of Coordination (OMC) for the Research and Innovation Policies in the Building of the European Research Area (ERA)

The Open method of coordination (MOC) is a procedure by which the European Commission helps member states to have their policies converge. After the Lisbon European Council in 2000 it is starting to be applied to the Research and Innovation Policies in the context of the building of the European Research Area (ERA). As stated (Commission of the European Communities, 2003), there is a “need for a collective process of monitoring and reporting on national policies and initiatives”.

The method, which is largely based on benchmarking through comparative indicators, involves, amongst other elements: (a) fixing guidelines for the whole Union for achieving the goals; and (b) “establishing quantitative and qualitative indicators and benchmarks (…) as a means of comparing best practice” (Commission Services, 2000).

The European Commission had been asked to prepare a benchmarking report on the five following themes:

– Human resources in RTD;
– Public and private investment in RTD;
– Scientific and technological productivity;
– Impact of RTD on economic competitiveness and employment;
– Promotion of RTD culture and public understanding of science.

The Commission produced the figures for a set of indicators related to each of the themes (Table 5.2) (European Commission, Key figures, 2001) and established five expert groups to conduct the analysis of these themes, that have produced reports which offer a comprehensive review of the themes selected by the Research Council in June 2001 (European Commission, Benchmarking reports, 2001).

The Working Group on ‘Benchmarking and S&T Productivity’ (STRATA–ETAN Working Group, 2002a) explicitly endorses what Lundvall and Tomlinson (2001) have called ‘intelligent benchmarking’. This leads the group to state that the indicators mean different things in different countries, which are characterised by systemic differences and situations that are context specific and path–dependent. The notion of ‘learning by comparing’ is then put forward. The consequence is that ‘quantitative comparison (…) is only the starting point for further analysis’.

In the same vein, the Working Group on “The Impact of RTD on Competitiveness and Employment” (STRATA–ETAN Expert Group, 2002b), states that “the indicators are of most interest when they can be related to theoretical concepts concerning the functioning of the innovation systems", stressing the non-trivial and debatable aspects of the interpretation of an indicator. Similarly, the Working Group on 'Human Resources in RTD' (STRATA–ETAN Working Group, 2002c), point to the fact of "indicators (…) must be followed by more comprehensive studies of the area being benchmarked".

In those benchmarking exercises 'intelligent benchmarking' is called for, which is very much a process in which indicators, as analytical knowledge, do not close the debates, but on the contrary contribute to it by the criticism they raise or the diverging interpretation they generate. The role of S&T indicators in the OMC is a good example of this situation.

3.2 The European Innovation Scoreboard (EIS) of the EU

The European Innovation Scoreboard (EIS) has been developed by the European Commission as requested by the Lisbon Council in March 2000. It supports 'transnational policy learning' amongst policy makers in the area of innovation, and, in this perspective, it compiles a set of commented indicators under four categories (see Table 5.1):

– Human resources;
– Creation of new knowledge;
– Transmission and application of knowledge;
– Innovation finance, outputs and markets.

It is revealing that the EU officials present these indicators as "a starting point for discussion and action" (European Commission, Scoreboard, 2002). They go on to say that there are "economic and cultural reasons for high or low scores in individual indicators, and (that) some may criticise exactly what is measured (…) which gives deeper insight into why there are such worrying disparities."

Interestingly, benchmarking workshops are organised, each on a specific innovation policy theme. Discussions at these workshops go into depth on selected national policies and schemes, providing policy makers the opportunity to embark on 'intelligent benchmarking', taking into account the diversity of approaches and the differences in national context.

The Evaluation Report of the Trendchart Policy Benchmarking workshops 2001/2002 (July 2003) (European Commission, Trendchart Report, 2002) assesses that one of the impacts of the workshops is that the "discussion on indicators is used in further discussion".

3.3 Benchmarking Industry–Science Relationships: the OECD Pilot Study on France and the UK

As part of the OECD project on benchmarking industry–science relationships (ISRs), the objective of the pilot study was to experiment with benchmarking as a tool for policy diagnosis and to contribute to the development of a benchmarking methodology suitable for public policy purposes (OECD, 2002). The benchmarking process (Figure 5.1) involved a series of seminars, fed by indicators and methodological papers.

Table 5.1. Indicators for the benchmarking of national research policies

Human resources in RTD

1. Number of researchers in relation to the total workforce
2. Number of new S&T PhDs in relation to the population in corresponding age group
3. Number of young researchers recruited in universities and public research centres
4. Proportion of women in the total number of researchers in universities and public research centres
5. Proportion of researchers from other countries amongst researchers in universities

Public and private investment in RTD

6. Total RD expenditure in relation to GDP and breakdown by source of funding
7. Research and development expenditure financed by industry in relation to industrial output
8. Share of the annual government budget allocated to research
9. Share of SMEs in publicly funded RD executed by the business sector
10. Volume of venture capital investment in early stages (seed and start-up) in relation to GDP

Scientific and technological productivity

11. Number of patents at the European and US patent offices per capita
12. Number of scientific publications and most cited publications per capita
13. Number of spin-offs generated by universities and research centres
14. Percentage of innovative firms cooperating with other firms/universities/public research institutes
15. Rate of usage of broadband electronic networks for research by RD laboratories.

Impact of RTD on economic competitiveness and employment

16. Growth rate of labour productivity
17. Share of high tech and medium-high tech industries in total employment and output
18. Share of knowledge intensive services in total employment & output
19. Technology balance of payments receipts as a proportion of GDP
20. Growth in a country’s world market share of exports of high tech products

RTD: research & technological development. RD: research & development SME: small and medium size enterprises. GDP: gross domestic product

Table 5.2. Indicators of the European Innovation Scoreboard

Human resources	
1.1	S&E graduates (‰ of 20–29 years age class)
1.2	Population with tertiary education (% of 25–64 years age class)
1.3	Participation in life long learning (% of 25–64 years age class)
1.4	Employment in medium high and high tech manufacturing (% of total workforce)
1.5	Employment in high tech services (% of total workforce)
Knowledge creation	
2.1	Public RD expenditures (% of GDP)
2.2	Business expenditures on RD (% of GDP)
2.3.1	EPO high tech patent applications (per million population)
2.3.2	USPTO high tech patent applications (per million population)
2.4.1	EPO patent applications (per million population)
2.4.2	USPTO patents granted (per million population)
Transmission and application of knowledge	
3.1	SMEs innovating in house (% of manufacturing SMEs and % of services SMEs)
3.2	SMEs involved in innovation co-operation (% of manufacturing SMEs and % of services SMEs)
3.3	Innovation expenditures (% of all turnover in manufacturing and % of all turnover in services)
Innovation finance, output and market	
4.1	Share of high tech venture capital investment
4.2	Share of early stage venture capital in GDP
4.3.1	SMEs sales of ‘new to market’ products (% of all turnover in manufacturing SMEs and % of all turnover in services SMEs)
4.3.2	SME sales of ‘new to the firm but not new to the market' products (% of all turnover in manufacturing SMEs and % of all turnover in services SMEs)
4.4	Internet access/use
4.5	Information and communication technologies expenditures (% of GDP)
4.6	Share of manufacturing value–added in high tech sectors
4.7	Volatility rates of SMEs (% of manufacturing SMEs and % of services

EPO: European patent office.
USPTO: United States patent and trademark office.
SME: small and medium size enterprises.
GDP: gross domestic product.

The following highlights the results of the pilot study regarding benchmarking methodology for policy.

a) Avoiding two traps

The first trap is to see benchmarking as simple arithmetic on a few basic indicators. The assumption is that industry–science relationships (ISRs) can be characterised by a few simple input and output indicators, and ranked according to measured performance. This approach has the advantage of producing results which are easy to communicate to decision makers. However, its cognitive content is extremely weak and it can lead to wrong policy recommendations. In addition, it does not create incentives for improvement since the actors do not recognise themselves in such simplistic procedure.

The second trap is to seek for depth at the expense of international comparability, and consequently be overwhelmed by countries' idiosyncrasies. Benchmarking is reduced to a set of case studies which can only be put side by side. Whilst this approach can mobilise actors and provide interesting information, it does not challenge current practices and thus creates little 'peer pressure' on policy makers.

b) Benchmarking as a joint learning and strategic evaluation process

The challenge for the pilot study was to find a way around these traps. This was done through a structured and iterative process of interaction amongst stakeholders from the two countries in addressing common issues, based on both indicators (quantification — codified knowledge) and qualitative information and expert judgement (contextualisation — tacit knowledge).

The sequence of activities is the following (Figure 5.1), the basic idea being to combine qualitative analysis, quantitative indicators, and expert judgements within a common conceptual framework so as to ensure as much comparability as possible:

- Establishment of a conceptual framework providing classifications of linkages, institutional arrangements, and incentives structure;
- Preparation of a background report (a) to relate the conceptual framework's categories to the 'real' national innovation system, its institutions and actors, and (b) to present internationally comparable quantitative indicators of the intensity and quality of ISRs;
- Working of an expert group in three benchmarking seminars: (a) interpreting the quantitative indicators accounting for national specificity of various kinds; (b) identifying of the key determinants of ISRs' effectiveness, as well as the strengths and weaknesses of each national system of ISRs; and (c) relating these determinants (strengths/

weaknesses) to policy options in key areas of government policy with a view.

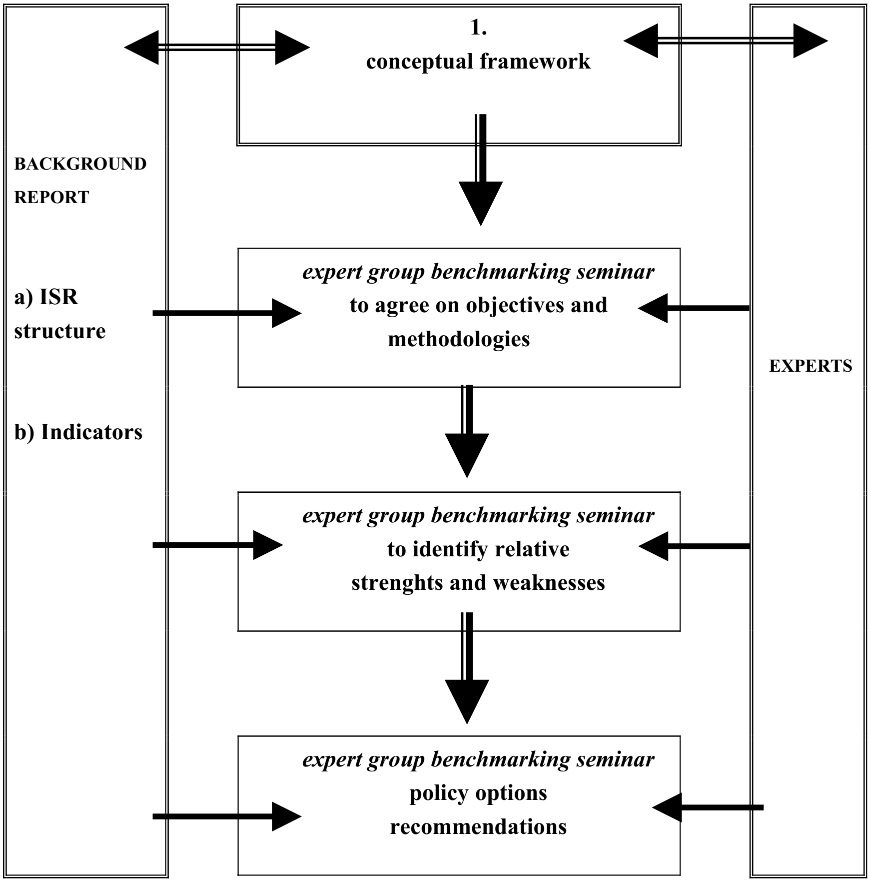


Figure 5.1: The proposed overall benchmarking process

3.4 France and United Kingdom Public Research Productivity Debate

In an article published by *Science* Sir Robert May (1998), at the time Chief Scientist of the British Government, proposed an 'indicator of efficiency of public research spending', constructed in the following way:

Number of scientific[1] publications / spending on basic research

In 1996 the value of the indicator, as computed by Sir Robert May, was 2.13 times higher for the United Kingdom than for France[2]. His conclusion was that public research spending was 2.13 times more efficient in the UK than in France. The case included all the ingredients of an apparently perfect case of benchmarking scientific productivity: an explicitly defined indicator of productivity, an international comparison, and a policy context at the highest level.

In view of what could be considered a flawed use of S&T indicators, an article was produced to show that the conclusion drawn was in essence illegitimate and that the ratio of 'number of publications' to 'spending' could not be interpreted as representing the efficiency of public research spending (Barré, 2001b).

The idea is first to discuss which figures are put, in each country, under the heading 'basic research' and second to control for the bias in the world share of articles introduced by the strong specialisation on the UK in clinical medicine. Then other possible biases are controlled, such as differences in salary and social benefits amongst the two countries, the weight of the social sciences and humanities, the linguistic differences, publications resulting from non-public spending. At the end, it was shown that those possible biases had the same order of magnitude as the gap initially labelled "differences in productivity".

It finally appears that productivity cannot be characterised by the ratio of one output variable (amongst many) to one input variable (amongst many) (ETAN Expert Working Group, 1999; Georghiou, 1999).

Indeed, that the productivity of a research system cannot be captured by a single number directly derives from what we have learned over the past 20 years about research activities: that they involve systemic and interactive dynamics among a variety of actors, linked to institutions, themselves embedded in society.

[1] Scientific publications recorded in the *Science Citation Index* (SCI), a data base built from about 3,500 Scientific journals which are the most cited. In what follows we shall refer to these Scientific journals as 'SCI journals'.

[2] The UK publishes 18.83 articles per million GBP and France 8.86.

4. IMPLICATIONS OF THE MODEL FOR THE WORK OF BIBLIOMETRICIANS AND S&T INDICATORS SPECIALISTS

What results from the above is that any mechanistic direct linkage between indicators and assessment for policy purposes must be disregarded. This introduces the idea of an informed pluralistic debate in a wider network, involving decision makers. It suggests that quantitative indicators are useful, but only as entry points into the discussion, considering their *raison d'être* as being to be criticised in terms of their (limited) relevance and (limited) comparability. It also suggests that such informed criticism provides an excellent way to deepen the understanding of the mechanisms at play in a comparative way.

We argue that this has implications on the work and responsibilities of the S&T indicators specialists: not only must they build reliable and relevant indicators, which are the classical criteria, but they must also add another criterion to the fulfilment of their duties, which we label the *Agora process* criterion.

4.1 The Reliability[3] Criterion

The reliability of an indicator is the confidence one can have that it measures exactly what it pretends to measure. There are two components to reliability:

– The accuracy of the computation of the indicator;
– The coherence between what is measured and what it is supposed to measure.

The accuracy of computation is obtained through the transparency of the data collection and treatment processes. Source data must be referenced, treatments must be made explicit and reproducible. The production of the indicators has to be challengeable and disputable. The Popperian criterion of what is a scientific knowledge applies: there has to be a procedure which can prove it is false.

In practice, limitations on accuracy come from difficulties in having full transparency in the production of source data and the complexity of the

[3] The methodological literature uses the term 'validity'; we use, nevertheless, the term 'reliability' which conveys a more specific meaning for the non-specialist in methodology than the term validity.

treatments, which make it almost impossible to make explicit and share the various steps. Some sort of validation of computation occurs through comparison of results obtained by different indicators producing units, hence the importance of the diversity of indicators production capabilities.

The coherence between what is measured in reality and what it is supposed to measure depends on the conceptual aspects of production of indicators. Two questions can be mentioned:

– How precise are the definitions of the measured parameters? For example, there are many ways of defining what is a researcher (one can include, or not, the post doctoral students), so that risks of non-comparability are high between two measurements, unless it is very precisely indicated what is the definition used.
– What is the accuracy of the correspondence between the parameter actually measured and its supposed meaning? This is the whole problem of the 'proxies': for example, a classical 'proxy' for measuring the technological orientation of a public research institution or a university, is to build an indicator of its patenting activity. But in the case in which the patenting is constrained by the resources devoted to the patenting bureau of the institution or university, the proxy will measure the evolution of the budget of the bureau, and not the technological capability of the institution.

4.2 The Relevance Criterion

In the definition of an indicator its relevance for decision making processes is of particular importance. S&T indicators have to address the questions which are at stake in all possible contexts. They need to measure parameters of entities of various scales, from laboratory to national level (micro, meso, macro scales) and parameters describing the human and financial resources aspects (inputs), the S&T production aspects (outputs) or the institutional, financial or cognitive interactions (co-operations, linkages, knowledge flows).

In brief, the relevance of an indicator will depend on:

– The proper understanding of what is at stake and of what is the need of the stakeholders and decision makers;
– The quality of the underlying conceptual model, which helps define both the parameters to be measured and the hypotheses to be tested and discussed.

4.3 The *Agora Process* Criterion

This additional criterion refers to the need for indicator specialists to be part of the workshops and debates where the indicators are interpreted and criticized by the stakeholders in their attempt to assess the situation and the policy to be recommended.

The indicator specialists will not have a purely technical role, but will have to understand the arguments of the stakeholders and to be able to link them to the indicators. He/she will have the role of a mediator between the stakeholders: beyond the problems of bias and proxies, the interpretation of an indicator rests on hypotheses which are not necessarily shared by all actors.

Transparency of the production of S&T indicators, making explicit the underlying concepts, identification of approximations, argumentation of the validity of the proxies become part of the S&T indicators building mandate.

This points towards excellence and three aspects simultaneously:

- The technical aspects of building indicators;
- The knowledge of the theoretical concepts underlying the indicators, that is the capability of understanding the indicators as proxies for concepts or parameters taking their significance in innovation theory;
- The *Agora processes* aspects, which call for the capability of going along with the indicators in the place where they are discussed, because the indicators specialist has an irreplaceable role in bringing the relevant background information for the debate based on the indicators to be rich and relevant.

5. CONCLUSION

S&T indicators become literally the common language allowing the stakeholders to be part of the exploration and assessment of the technological trajectories which shape innovation, allowing for a transparent, credible, and acceptable decision processes — in other words addressing concretely the question of the democratic dimension of science, technology and innovation.

In this model S&T indicators lose their position of being unquestionable, but gain a central position as a link between the actors (research organisations, policy makers, the media, the political circles): S&T indicators appear to be at centre of stage as a device having unique capabilities of allowing for inter-operability of the visions and discourses of the stakeholders. They become the standard intermediary through which governments assess their policies and practices, after due debate and exchanges on the possible interpretations of the figures.

It is for the new generation of S&T indicators specialists to face the double challenge first of producing relevant and sound indicators and second to go with their indicators along the debates which they have to make possible.

Only then will S&T indicators fulfil their potential as instruments of objectivation, mediation, and participation, so crucially needed in an era of policy making of changing science–society relationships.

REFERENCES

Barré, R. (2001a). The Agora model of innovation systems: S&T indicators for a democratic knowledge society. *Research Evaluation*, 10 (1), 13–18.

Barré, R. (2001b). Sense and nonsense of S&T productivity indicators. *Science and Public Policy*, 28 (4), 1–8.

Barré, R. (2002). *S&T policy making fort the future: new rationales, new design tools*. In: S&T policies in Europe: new challenges and new responses. STRATA Consolidating Workshop, Brussels 22–23 April 2002, Report EUR 20440, (pp. 86–134).

Caracostas, P., Muldur, U. (1997). *Society, the Endless Frontier*. European Commission, EUR 17655, Brussels.

Commission of the European Communities (2003). *Investing in research: an action plan for Europe*. COM(2003)226, Brussels.

Commission services (2000). *Development of an open method of coordination for benchmarking national research policies — objectives, methodology and indicators*. SEC(2000)1842, Brussels.

European Commission (2001). *Key figures.* http://www.cordis.lu/indicators/pressconf_kfbd.htm.

European Commission (2001). Benchmarking reports. 2001 http://www.cordis.lu/era/benchmarking.htm.

European Commission (2002). Scoreboard. http://www.cordis.lu/ scoreboard/what.htm

European Commission (2002). Trendchart Report. http://trendchart.cordis.lu/Reports/Documents/workshop_evaluation_030701rw2.pdf.

Etzkowitz H., Leydesdorff, L. (2000). The dynamics of innovation: from National Systems and ‘Mode 2’ to a Triple Helix of University–Industry–Government Relations, introduction to the special “Triple Helix” issue*. Research Policy* 29 (2), 109–123.

Georghiou, L. (1999). Meta–evaluation: evaluation of evaluations. Special Issue on Proceedings of the Conference Science and the Academic System in Transition — An International Expert Meeting on Evaluation — Vienna, July 3–5, *Scientometrics*, 45, 523–530.

Kuhlmann, S., et al. (1999). *Improving distributed intelligence in complex innovation systems.* ASTPP network — TSER Contract SOE1-CT96-1013, Karlsruhe.

Kuhlmann, S., Edler, J. (2003a). Scenarios of technology and innovation policies in Europe: Investigating Future Governance, *Technological Forecasting and Social Change* 70, 619–637.

Kuhlmann, S. (2003b). *Evaluation as a source of 'strategic intelligence'*. In Ph. Shapira, S. Kuhlmann (Eds.), Learning from Science and Technology Policy Evaluation: Experiences from the United States and Europe (pp.352–379). Cheltenham: E. Elgar.

Latour, B. (1999). *Politiques de la nature – comment faire entrer les sciences en démocratie*. La Découverte: Paris.

Lundvall B-A. and M. Tomlinson, (2001). *Learning by comparing: reflection on the use and abuse of benchmarking.* In G. Sweeney (Ed.), Innovation, economic progress and quality of life. London: Elgar Publishers.

May, R.M. (1998). The scientific investments of nations. *Science,* 281, 49–51.

Nowotny, N., Scott, P., Gibbons, M. (2001). *Re-thinking science; knowledge and the public in an age of uncertainty*. Polity Press: Cambridge, UK.

OECD (2002). *Benchmarking Industry–Science relationships*. Paris.

Smits, R. (2001). *The new role of strategic intelligence*. In A. Tubke, K. Ducatel, J. Gavigan, P. Moncada (Eds.), Strategic policy intelligence: current trends, the state of play and perspectives. JRC-ESTO, EUR 20137 EN 1–30, Brussels.

STRATA-ETAN working Group (2002a). *Benchmarking national RD policies: S&T productivity*. European Commission: Brussels (http://www.cordis.lu/era/benchmarking.htm).

STRATA-ETAN expert Group (2002b). *Benchmarking national research policies; the impact of RTD on competitiveness and employment (IRCE).* European Commission: Brussels (http://www.cordis.lu/era/benchmarking.htm).

STRATA-ETAN working Group (2002c). *Benchmarking national RD policies: Human resources in RTD*. European Commission: Brussels (http://www.cordis.lu/era/benchmarking.htm).

Chapter 6

PARADIGMS AND TRAJECTORIES OF TECHNOLOGICAL OPPORTUNITIES 1890–1990

An Evolutionary Economics and Patent Statistics Approach

Birgitte Andersen

Birkbeck, University of London, UK. E-mail: b.andersen@bbk.ac.uk

Abstract: This chapter examines the evolution of technological opportunities historically. The disciplinary approach is evolutionary economics combined with patent statistics. By analysing the complexities behind changing technological opportunities, technological paradigms governing the evolutionary process of change are traced empirically. Evidence shows how the path-dependent evolution of technological opportunities is characterised by 'creative, incremental, accumulation', as it is revealed how new technological systems with new opportunities build upon (complement and extend) the knowledge embodied in old ones, rather than substitute them. Evidence also reveals how the evolution of technological opportunities has become increasingly interrelated, wider-ranging and complex, in which trajectories previously following isolated channels of development are brought together. Finally, evidence shows how typical technological trajectories of broad technological groups explain technological evolution better than the conventional aggregate measures of such broad technological fields.

1. INTRODUCTION

The major source of inspiration in this chapter is provided by the Schumpeterian evolutionary economic approach to institutional economics. It has been argued by institutional evolutionary economists that changing technological opportunities along trajectories governed by technological paradigms or regimes is perhaps the most central regulating variable in society (Dosi, 1982, 1988; Perez, 1983; Freeman, Clark and Soete, 1982; Freeman and Perez, 1988). Basically, when evolutionary institutional economists point to the existence of technological trajectories some

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 133-161.

trajectories are more likely to be followed than others, as defined by the existing paradigm. There has to be distinguished between those trajectories that are specific to a particular technology, product or industry, and those that are of general importance (Nelson and Winter, 1977). Freeman and Perez (1988) differ from Dosi (1988) in the sense that they refer to the Schumpeterian type of *meta*-paradigm of a dominant technological (or techno-economic) regime which rules for several decades, whereas Dosi refers to a type of (micro) paradigm or trajectory that is specific to a particular technology. In Freeman and Perez's framework a radical change in the whole economy is related to a *generalised* shift in the technological (or techno-economic) paradigm, revealing an overall shift in structures at the micro level as well as the macro level throughout the economic system. Accordingly, Freeman and Perez refer to generalised structural changes as the hard core of long term theorising. This way of theorising can also, somewhat, be compared to Kuhn's (1962) notion of scientific revolutions.

Furthermore, evolutionary institutional economists have agreed that, owing to the importance of understanding the impact of technology on the dynamics of the economy, theory must pay special attention to the origin and history of new technologies as technological development evolves accumulative, incremental and path-dependent (Rosenberg, 1976, Nelson and Winter, 1977, 1982). However, any attempts to identify quantitatively, and statistically measure, the nature of the evolution of technological opportunities has lagged behind more qualitative accounts.

Thus, this chapter aims to identify quantitatively, and statistically measure, the changes in technological opportunities during the last century of technological evolution (1890–1990), in order to subsequently trace empirically the evolution of 'trajectories' of technological opportunities governed by 'technological paradigms' or regimes.

The emphasis in this chapter will only be on inventions and innovations that have diffused and led to generalised technological and economic changes, so this chapter operates within the framework of *meta*-paradigms. It is argued here that *ex post* the successful areas of new technological opportunities, or perhaps more correctly, those which hold up, reflect not just the technological opportunities which have governed and been governed by the paradigms, but also the areas in which society possessed socio-economic competence, because without that the paradigm would not have been unleashed. Hence, although the focus is narrowed to the technological features of the evolution of paradigms, the characteristics by technological epochs, divided by *overall* structural changes in the evolution pattern of technological opportunities, can be regarded as a reflection of the overall features of the paradigm. The specific development paths of opportunities for selected technological sectors are in this context defined as technological trajectories.

Furthermore, it is assumed here that a fast rate of growth of patenting (i.e. a high rate of growth in patent stock) in a technological sector of activity represents an area of strong technological opportunity in the period in question (see section 2). This is now a standard assumption within patent statistics in general.

The chapter is structured as follows: The selection and organisation of the patent data on which this chapter is based will first be presented. This includes patents belonging to Chemical, Electrical/electronics, Mechanical, Transport, and Non-industrial sectors of technological activity. After that, the areas of greatest technological opportunities within different epochs of technological development will then be calculated and subsequently used to extract the complexities behind the changing technological opportunities within and between broad technological groups historically. The purpose of this is first to sketch technological epochs of *overall* structural changes in patenting patterns governed by the evolution of technological paradigms. Then there will be an investigation of the evolving nature of technological opportunities. It will be investigated whether the composition of technological opportunities changes gradually (i.e. evolves relatively stable), or whether the changes are characterised by radical structural disruptions across historical epochs of development. The extent to which evolution is led by divergence or convergence in composition of technological opportunities historically will also be the subject for investigation. This will enable an examination of the extent to which paradigms governing new technological epochs ‘creatively’ destroy old ones or complement and extend them. Finally, different possible trajectories of technological opportunities will be set out, and each broad technological group’s relative contribution to specific paths will be calculated. This will enable us to identify typical (or revealed) technological trajectories of great importance for each broad technological group.

It should be noted that after the data are presented and graphically illustrated, the data on Non-industrial patent classes or sectors will *only* be included in the analysis when the inter-class (or inter-sector) relative distribution for all sectors across all broad groups (i.e. for total) plays a role. However, it will not make sense to conduct any analyses at the Non-industrial group level, because the group is a residual group and therefore characterised by great technological heterogeneity, as opposed to the Chemical, Electrical/electronics, Mechanical, and Transport broad technological groups.

2. THE DATA

This paper is based on a US patent database which has been constructed by Professor John Cantwell with the assistance of the US Patent and Trademark Office. The database comprises both individual and corporate patents granted in the US from 1890 to 1990. Each patent is classified by the year in which it was granted and by the type of technological activity with which it is most associated. In this context, patent classes (or a sub-division of a class) have been allocated to one of 399 technological sectors, which in turn belong to one of 5 broad technological groups, consisting of Chemical, Electrical/electronic, Mechanical and Transport technologies, plus a residual consisting of other mainly Non-industrial technologies. The types of technologies which are most characteristic for the broad technological groups are presented in detail in Andersen (2001, chapter 2), which also addresses relevant issues related to the construction of the generously proportioned historical patent data set. However, the sectoral classification of patents according to the type of technological activity with which each patent is associated must be distinguished from the industry of the firms to which patents may be assigned, both of which have been recorded separately. Most large firms have engaged in at least some development in most of the broad technological groups of activity, irrespective of the industry in which they operate (see Andersen 2001, chapters 6 and 7).

Because a patent has to reflect a novelty (i.e. a movement of the technological frontier) it has often been used as an economic indicator when measuring the rate and direction of technological change and corporate innovation. This chapter will use patent statistics to contribute to the discussion concerning determining the rates and directions of technological change in the evolution of technological activities and opportunities. Thus patent data here serve as an indicator for two variables: (i) 'Accumulated technological impact or socio-economic importance' which is acquired from innovation over time, and hence reflected in the stock of a patent class or technological sector. (ii) 'The extent of technological opportunity' which is reflected in the growth rate of the patent stock of the class or sector in question. As patent data is only a direct measure of invention, there are equally some *potential difficulties* with this approach in which patent data serve as proxy measures. Andersen (2001, chapter 2) critically reviews those potential difficulties and explains some statistical means for accounting for or minimising such difficulties.

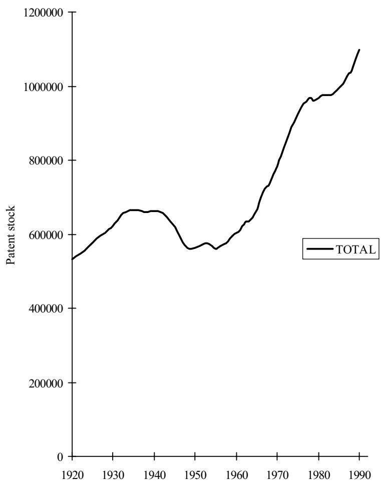
There are many advantages of working with patent stocks, and it is also now commonly recognised that the most serious drawbacks that are involved in the (inappropriate) use of patent statistics — most notably, stochastic fluctuations in variations in the importance of individual patents (and in the propensity to patent across sectors and over time) — are substantially ameliorated over large numbers of patents. Other advantages of using patent stocks and growth rate in patent stocks in patent statistics are reviewed in Andersen (2001, chapter 2)

Cumulative stocks of patents have been calculated for technological sectors, technological groups, broader technological groups, and total. Stocks were calculated using the perpetual inventory method as in vintage capital models, with an allowance for a depreciation of the separate contribution of each new item of technological knowledge over a thirty year period. This is the normal assumption for the average lifetime of capital, given that new technological knowledge is partly embodied in new equipment or devices, as well as the conventional method used in patent statistics, as introduced by Cantwell and his associates who pioneered the use of accumulated stocks of long time series within patent statistics. (Previous work on long time series with patents was, of course, done by Kuznets (1930) who accumulated stocks only over 5 year periods and investigated the change in those, and by Schmookler (1966) who worked with flows.) Thus the stock in 1919 represents a weighted accumulation of patenting between 1890 and 1919, a thirty year period with weights rising on a linear scale from 1/30 in 1890 to unity (30/30) in 1919, using ‘straight line depreciation’. Although the assumption of a thirty-year life is admittedly arbitrary, it must be emphasised that this is a proxy measure of the life of the *underlying technological knowledge* and the tangible devices with which it is associated. Hence it is not intended to capture the lifetime of the patent (which is shorter) or the life of the economic value of the patent. However, the results would be largely unaffected if a shorter lifetime was assumed. Although patent stocks would then fluctuate more as the smoothing process associated with accumulation would be less pronounced, and so the absolute values of the growth of stocks would be greater, the identification of the periods in which stocks grow relatively faster or slower would be largely unaffected.

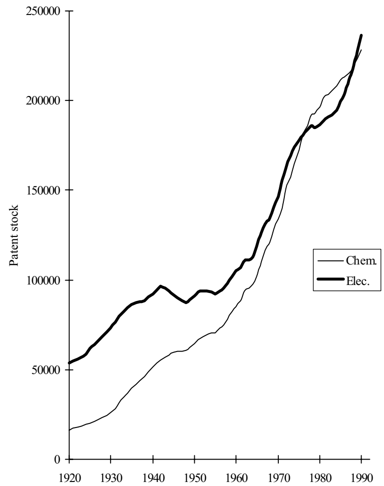
3. TRACING TECHNOLOGICAL PARADIGMS

To get some ideas of the historical trends at the macro level accumulated patent stocks, calculated in section 2, for the aggregate of all patent classes and for each of the five broad technological groups are displayed in Graphs 1 to 4 included in Figure 6.1. The graphs certainly suggest changes in technological opportunities at the aggregate level over time.

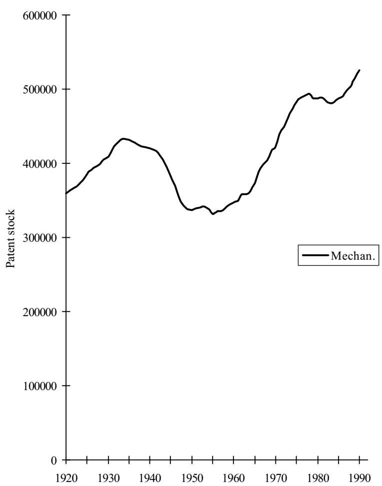
Graph 1. TOTAL accumulated patent stock 1920-1990



Graph 2. Accumulated patent stocks 1920-1990 for CHEMICAL and **ELECTRICAL/ ELECTRONIC** technologies



Graph 3. Accumulated patent stock 1920-1990 for MECHANICAL technologies



Graph 4. Accumulated patent stocks 1920-1990 for **TRANSPORT** and NON-INDUSTRIAL technologies

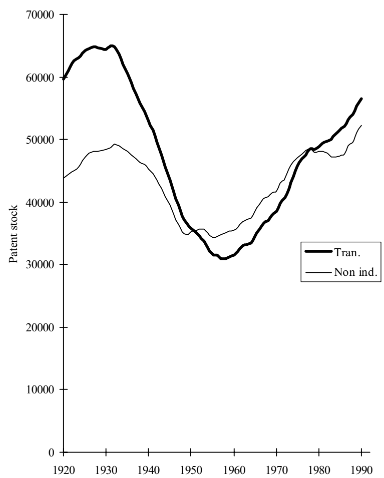


Figure 6.1. Accumulated patent stocks 1920–1990

It can be seen that the growing opportunities in the science based sectors (Chemical and Electrical/electronics) have been more or less continuous in the twentieth century, except for some disruptions in the growth rates between 1940 and 1960; whilst the opportunities in the engineering–based sectors (Mechanical and Transport) as well as the Non-industrial group were weak between the 1930s and up to 1960, over which period those broad technological groups even experienced an actual decline in accumulated patent stock.

The combination of these effects suggest that the appropriate time periods into which to split the analysis are 1920–40 (the Interwar period), 1940–60 (the War/early postwar period) and 1960–90 (the Recent period), as the general picture shows interruptions or breaks in the trends between these periods. The wars referred to in this context are World War I and World War II. However, as economists within the evolutionary tradition emphasise, the macro economy is not simply the aggregate of various micro units but is instead regarded as a complex outcome of micro relationships or interactions, the further analysis will be carried out at a more disaggregated level in order to understand the evolving structures which lie behind the shapes of the aggregate graphs.

When examining the complexities behind the structure of the changing technological opportunities, only patent classes which have at least 10 patents in the beginning and at the end of each of the selected periods have been included, in order to avoid problems which can be created by small numbers, which can lead to very high positive or negative growth rates and which represent only random statistical results. Hence, out of the 399 patent classes 369 satisfy the selection criteria, of which the Chemical broad group includes 50 (down by 7), Electrical/electronics accounts for 57 (down by 12), Mechanical for 213 (down by 8), Transport for 21 (no change) and the Non-industrial group for 28 (down by 3).

The distribution of the total number of the 369 selected patent classes ranked in accordance with their technological opportunities or growth rates (high, medium, or low), whether in absolute or relative terms, and over the three historical periods (i.e. interwar 1920–40; war/early postwar 1940–60; and recent period 1960–90), is presented in Table 6.1.

From Table 6.1 it can be observed that the areas of greatest technological opportunities are not concentrated within relatively few areas of presumably related technological fields, but have been increasingly widely dispersed across the five broad technological groups.

Table 6.1. Classes within five broad technological groups; ranked in accordance with their technological opportunities or growth rates over three historical periods

Growth rate rankings		*Number of sectors*			*Intra-'broad technological group' distribution in %*		
	Period →	1920-1940	1940-1960	1960-1990	1920-1940	1940-1960	1960-1990
High	Chemicals	48	40	32	96.00	80.00	64.00
	Electrical/ electronics	33	31	27	57.89	54.39	47.37
	Mechanical	38	41	48	17.84	19.25	22.54
	Transport	2	1	9	9.52	4.76	42.86
	Non-industrial	2	10	7	7.14	35.71	25.00
	TOTAL	*123*	*123*	*123*	-	-	-
Medium	Chemicals	2	9	14	4.00	18.00	28.00
	Electrical/ electronics	17	16	15	29.82	28.07	26.32
	Mechanical	91	88	78	42.72	41.31	36.62
	Transport	4	4	3	19.05	19.05	14.29
	Non-industrial	9	6	13	32.14	21.43	46.43
	TOTAL	*123*	*123*	*123*	-	-	-
Low	Chemicals	0	1	4	0.00	2.00	8.00
	Electrical/ electronics	7	10	15	12.28	17.54	26.32
	Mechanical	84	84	87	39.44	39.44	40.85
	Transport	15	16	9	71.43	76.19	42.86
	Non-industrial	17	12	8	60.71	2.86	28.57
	TOTAL	*123*	*123*	*123*	-	-	-

In this context Chemicals and Electrical/electronics (the science–based sectors), which at the beginning of the 20th century had a relatively high proportion of sectors ranked amongst the fastest growing, have generally seen a decline in their share of the number of technological sectors ranked among the fastest growing, whilst Mechanical, Transport (the engineering–based sectors) and Non-industrial, which at the beginning of this century had a low proportion of sectors ranked amongst the fastest growing, have generally seen an increase in their share of the number of the technological sectors ranked amongst the fastest growing. However, the greater fluctuations in the relative growth rate rankings of patenting in the Transport and Non-industrial spheres can be partly explained by the relatively small total number of patent classes in those groups in comparison with the other broad technological groups. Two alternative interpretations concerning changes in the composition of the band with fastest growing technologies (or areas with highest technological opportunities) might be provided, and they are classified into two models, Model A and Model B.

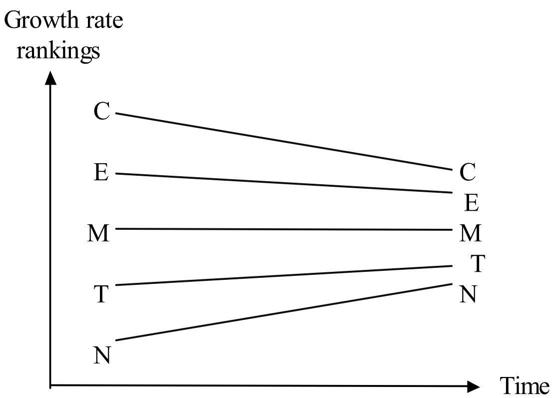


Figure 6.2. Illustration of inter-group convergence according to Model A

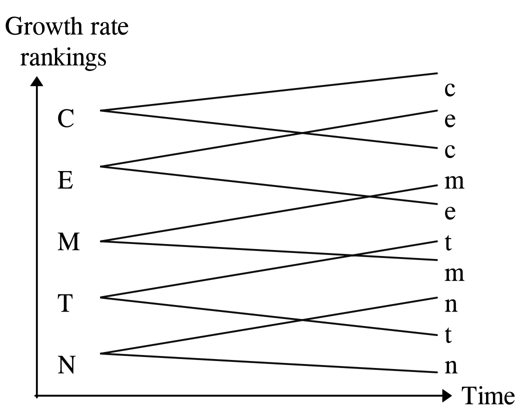


Figure 6.3. Illustration of intra-group dispersion according to Model B

Model A argues that the technological opportunities have become less concentrated or more interrelated as a result of inter-group convergence in technological opportunities (see Figure 6.2). The alternative explanation comes from Model B, which contends that the technological opportunities have become less concentrated owing to intra-group dispersion of technological opportunities (see Figure 6.3). The full description of predictions based on Model A and Model B is described in Tables 6.2 and 6.3.

Table 6.2. Predictions based on INTER-group dispersion (or concentration) of technological opportunities

Cross-technological group growth comparisons	*Greater INTER-group dispersion between periods*	*Greater INTER-group convergence between periods*
High growth rate ranked groups: (Chemicals and Electrical/electronics)	More in top ranked classes	Less in top ranked classes
Medium and low growth rate ranked groups: (Mechanical, Transport and Non-industrial)	Less in top ranked classes	More in top ranked classes

Model A describes how the INTER-group changes in the rankings of patent classes over time might have been associated with INTER-group dispersion (or concentration) in the growth of different broad groups of related technological fields.

Table 6.3. Predictions based on INTRA-group dispersion (or concentration) of technological opportunities

Cross-technological sector growth comparisons →	*Greater INTRA-group dispersion between periods*	*Greater INTRA-group convergence between periods*
Cross-technological group growth comparisons		
High growth rate ranked groups: (Chemicals and Electrical/electronics)	Less in top ranked classes	More in top ranked classes
Medium and low growth rate ranked groups: (Mechanical, Transport and Non-industrial)	More in top ranked classes	Less in top ranked classes

Model B describes how the INTER-group changes in the rankings of patent classes over time might have been associated with INTRA-group dispersion (or concentration) in the growth of different related technological fields collected together in each broad group.

Statistical evidence has been compiled to examine whether either Model A or Model B (or in some cases both of them) account for the trends described above. As the mean growth rates of the five technological groups vary between groups and over time (see Table 6.4), it is difficult to directly compare the degree of concentration and dispersion across classes within groups using the absolute measures provided by the variance or the standard deviation of growth rates, as these depend upon the variate scale. Therefore this analysis needs a relative measure of the variability of the growth rates across classes in the data which for example, is provided by the coefficient of variation, represented by the standard deviation divided by the mean expressed as a percentage:

(i) $CV = (\sigma/\mu) * 100\%$.

(From Table 6.4 it can also be observed that the results concerning convergence versus divergence of growth rates in most cases varies, depending on whether the standard deviation or the coefficient of variation has been used, and that the coefficient of variation is a better measure of concentration owing to the great importance of the moving average growth rates over time.)

For Model A the CV has been calculated using the unweighted average of all the growth rates of all the 369 individual patent classes; whilst for Model B the CV has been calculated across individual classes' growth rates within each technological group in each period using the unweighted average of all the growth rates within each group. The results of the average growth rates and CVs are displayed in Table 6.4.

In that way the models which explain best the complexities behind the evolution of technological opportunities have been found for all broad technological groups (except for the Non-industrial group, for the obvious reasons stated in see section 1). The results are listed in Table 6.5.

By comparing the statistical results of Model A and Model B (by viewing the changes in the CVs displayed in Table 6.4) with the actual general pattern of changes in the composition of the fastest growing classes (as described in Table 6.1) in relation to the model predictions expressed in Tables 6.2 and 6.3, the model can be found which best explains the shifts in the rankings of patent classes.

To use the Chemical broad technological group as an example, we see from Table 6.4 that we have inter-group divergence in technological growth rates between the interwar and the war/early postwar period as the CV for total rises. In accordance with Model A (which deals with inter-group issues, see Table 6.2) this means that more patent classes within Chemicals would be ranked amongst the fastest growing. However, from Table 6.1 we see that the Chemical technological group actually declines in share of its patent classes that is ranked among the fastest growing. Hence model A does not explain the observed empirical evolution.

Table 6.4. Results concerning the relevance of Model A and Model B in relation to five broad technological groups as well as total*: Technological growth rates (σ) and CVs

		1920–1940	*1940–1960*	*1960–1990*
Average growth rates: μ (expressed in %)	$\mu_{\text{TOTAL 369}}$	132.77	15.15	88.98
	μ_C (Chemicals)	364.48	104.91	200.83
	μ_E (Electrical/ electronics)	405.82	78.93	138.77
	μ_M (Mechanical)	34.36	-13.38	55.54
	μ_T (Transport)	0.65	-43.50	56.14
Model A	(σ_{369})	(405.02)	(114.68)	(147.34)
	CV_{369} **	305.05	756.94	165.58
Model B	(σ_C)	(299.76)	(122.21)	(217.07)
	$CV_{\text{Chemicals}}$ ***	82.24	116.49	108.09
	(σ_E)	(895.64)	(219.72)	(180.77)
	$CV_{\text{Electrical/electronicsm}}$ ****	220.70	278.36	130.26
	(σ_M)	(71.68)	(45.40)	(104.13)
	$CV_{\text{Mechanical}}$ *****	208.60	-339.24	187.50
	(σ_T)	(56.38)	(27.29)	(89.47)
	$CV_{\text{Transport}}$ ******	8673.84	-62.74	159.37

* Only the absolute value of the CV is important (hence the positive or negative sign in front of the CV is not relevant).
**$CV_{369} = (\sigma_{369} / \mu_{TOTAL\ 369}) \cdot 100\%$
***$CV_{Chemicals} = (\sigma_C / \mu_C) \cdot 100\%$
****$CV_{Electrical/electronics} = (\sigma_E / \mu_E) \cdot 100\%$
*****$CV_{Mechanical} = (\sigma_M / \mu_M) \cdot 100\%$
******$CV_{Transport} = (\sigma_T / \mu_T) \cdot 100\%$

However, from Table 6.4 it can also be observed that the CV for Chemicals increases between the interwar and the war/early postwar period, reflecting intra-group divergence in growth rates. This would, in accordance with Model B (which deals with intra-group issues; see Figure 6.3 and Table 6.3), cause fewer chemical technological classes to be ranked among the fastest growing. Because we see from Table 6.1 that the broad Chemical technological group actually declines in total number of classes ranked among the fastest growing, the changes in the growth rate rankings of patent classes can be concluded to be explained by Model B in this case.

In very general terms these results show that Model B dominates between the interwar and the war/early postwar period, whilst Model A dominates in the more recent period (see Table 6.5). Hence we find two main technological paradigms governing the last century.

Table 6.5. Best model to explain the evolution of technological opportunities

Technological groups	*Period: 1920–40 to 1940–60*	*Period: 1940–60 to 1960–90*
Chemicals	B	A
Electrical/electronics	B	A
Mechanical	B	A
Transport	A and/or B	A and/or B

The technological epoch of intra-group divergence in the three biggest broad technological groups (Chemicals, Electrical/electronics and Mechanical) between the first two periods may reflect the formation of specialised engineering and science based fields, and the period in which the leading sectors of each of these technological groups came to maturity and established the structure for which they are known today. For Transport technologies instead there was intra-group convergence between the interwar and the war/early postwar period; this effect, which reduced the number of fast growing Transport classes, was further reinforced by inter-group divergence, as represented by an increased CV across all technological fields and illustrated in Model A.

However, the period between the war/early postwar and up to the recent period demonstrates a technological epoch of inter-group technological convergence characterising the evolution of all broad technological groups. This may reflect a new paradigm of formation of broader technological systems as well as the development of complex technologies which are offshoots of the incremental nature of technological development. These results can also be interpreted along with von Tunzelmann's notion of growing technological complexity (von Tunzelmann, 1995) and Kodama's notion of technology fusion (Kodama, 1992). It is suggestive of a historical shift towards more integrated technological systems in recent times through the fusion of diverse and formerly separate branches of technology, which explain the closing of the growth rate gap between the science based technologies (Chemicals and Electrical/electronics) and the engineering based technologies (Mechanical and Transport). This evidence concerning closer connections between the principal technological families also supports the study by Patel and Pavitt (1994). In other words, technological development has become increasingly interrelated and complementary rather than independent and distinct.

That the evolution of Transport technologies up to recent times is, as the only broad technological group, characterised by Model B, showing dispersion of intra-group growth rates, might be related to the revival of certain transport technologies in recent times, as seen in Graph 4 in Figure 6.1.

However, what is interesting is not only the *overall* changes of structures, but to examine the nature by which the technological opportunities have evolved. Section 4 will investigate the nature by which the technological opportunities have changed their relative positions (i.e. disruption of structures) and changed in their distribution (i.e. convergence versus divergence), whereas section 5 will identify and examine individual trajectories of technological opportunities.

4. THE NATURE OF EVOLVING TECHNOLOGICAL OPPORTUNITIES

The above results concerning *overall* structural changes in paradigms of patenting patterns will now be studied in more detail, with reference to the nature of the changing compositions of specific technological sectors. The nature by which the technological opportunities have changed their relative positions and changed in their distribution will be the unit of analysis.

Within post-Schumpeterian approaches to evolutionary institutional economics it is often argued that the structure of technology, and therefore also technological opportunities (as well as most other variables) changes only gradually or incrementally over time (see, e.g., Rosenberg, 1976; Nelson and Winter, 1982), as opposed to radically. This argument has also been applied from biological analogies to economics, although the limitations of such analogies have been recognised (Hodgson, 1993). However, the direction(s) of these incremental changes is often unclear; such as, e.g., the extent that gradual or incremental changes cause convergence or divergence of technological activity.

These issues concerning (i) structural disruption (i.e. high mobility in the relative growth rate ranking of technological sectors) or incremental nature of changes, as well as (ii) whether such change causes convergence or divergence, can be measured empirically/statistically by examining the changing composition of technological growth rates across different historical periods.

Such analysis is better done at the intra-group level owing to the great difference in absolute size (i.e. number of technological fields or classifications) of the broad technological groups, which would eliminate the smaller broad technological groups' overall contribution to the results. This would display only Electrical/electronics and Mechanical major historical regimes rather than evolving structures. The analysis is based on the eligible patent classes for broad groups (Chemical, Electrical/electronics, Mechanical and Transport) selected in section 3. Again based on work in the previous section, appropriate time periods into which to split such analysis are 1920–40 (the interwar period), 1940–60 (the war/early postwar period), and 1960–90 (the recent period).

4.1 The Model

Although a statistical analysis, cross-section changes in the pattern of the patent growth rates are decomposed into a 'mobility effect' and a 'regression effect'. The 'mobility effect' tests for the degree of mobility in the growth rate rankings of patent sectors: big mobility or change in growth rate rankings (indicating disruption or radical changes of existing structures) versus small change in rankings (indicating an incremental nature) versus no change. The 'regression effect' tests for the direction of the evolution: convergence vs. divergence vs. no changes. The statistical principle is commonly known as a Galtonian regression.

In this framework the regression coefficient β measures the direction of the intra-group changes in the distributions of patent classes' relative growth rates over time (i.e. whether the intra-group growth rates tend to move closer to or further away from the mean). The magnitude of the 'regression effect' is measured by $(1 - \beta)$, because for $\beta = 1$ (or $1 - \beta = 0$) there is on average no intra-group convergence or divergence of growth rates.

As mentioned above, the other feature arising from the regression analysis is the simple test of the extent of mobility or fluctuations in growth rates across patent classes (i.e. whether the intra-group variations in growth rates indicate patterns of stability or structural disruption in ranking of technological activity or opportunities across sectors over time). In this framework the correlation coefficient ρ is a measure of the degree of mobility or disruption of patent classes' growth rate rankings over time. The magnitude of the 'mobility effect' is measured by $(1 - \rho)$, because for $\rho = 1$ (or $(1 - \rho = 0)$ there is no mobility.

However, it should be noted that results showing a regression towards the mean (indicating that the intra-group growth rates on average move closer towards the mean) shall not be confused with a convergence in the distribution of the growth rates (as calculated in section 3). Basically, we might be tempted to believe that because a regression towards the mean there has been a reduction in the distribution in the growth rates. But this is a fallacy (as also argued and illustrated with examples by Hart, 1994). If the mobility effect exceeds the regression effect (so $\beta > \rho$, even though $\beta < 1$)

then the variance of the distribution will rise. (This implies that the CV, as calculated in section 3, will rise too, unless the mean happens to rise by proportionally even more).

4.2 The Estimations

Two simple cross-section regressions of the technological sectors' growth rates expanding over three broad periods (from the interwar t - 2 to the war/early postwar t - 1; and from the war/early postwar t - 1 to the recent period t) are carried out. The analysis is carried out separately for each broad technological group, as mentioned above.

However, it is found that it is better to express the growth rates in logarithmic form (since the distribution of the size and hence growth is closer to a log normal than a normal distribution):

(ii) $\text{Log}\,(GT_i + 1) = \text{Log}\,P_{i(t)} - \text{Log}\,P_{i(t-1)}$

where GT denotes the rate of growth, and P denotes the patent stock of sector i in time t (or t-1).

Hence, the changing composition of the technological opportunities can then be statistically tested using following cross-section regressions:

Regression 1. Between the interwar (1920–40) and the war/early postwar period (1940–60)

(iii) $[\text{Log}\,(GT + 1)]_{i(t-1)} = \alpha + ß\,[\text{Log}\,(GT + 1)]_{i(t-2)} + \varepsilon_{(t-1)}$

where [Log (GT + 1)] refers to the logarithm of rate of growth in patent stock i over the time period in question.

Regression 2. Between the war/early postwar (1940–60) and the recent period (1960–90)

(iv) $[\text{Log}\,(GT + 1)]_{i(t)} = \alpha + ß\,[\text{Log}\,(GT + 1)]_{i(t-1)} + \varepsilon_{(t)}$

where [Log (GT + 1)] refers to the logarithm of rate of growth in patent stock i over the time period in question.

4.3 The Results

The results are displayed in Tables 6.6 and 6.7. From Tables 6.6 and 6.7 it can be observed that especially the relatively newly established science–based groups (Chemicals and Electrical/electronics) have each experienced a strong regression towards the mean throughout the century, reflecting an integration of historical (or relatively old) and new fields of development. In order words the knowledge embodied in the old regime, operating between the interwar and war/early postwar period, is also embodied in the new regime which has been governing the period up to recent time.

Moreover, the science based sectors have also experienced some mobility or disruptions in the compositions of the intra-group growth rates throughout the history, and in recent times this even exceeded the convergence effect of the regression. This 'mobility effect' may suggest that the relatively newly established chemical industry is still going through some degree of 'search and selection' process of development.

Table 6.6. The extent of continuity (convergence or divergence) in the composition of technological opportunities within broad technological groups over time

Technological broad group	*Numbers of sectors*	*Period*	*Regression Effect* $\hat{\beta}$	$(1-\hat{\beta})$	$t_{\beta 1}$
Chemicals	50	Regression 1	0.450	0.550	-4.632***
	50	Regression 2	0.119	0.881	-5.371***
Electrical/ electronics	57	Regression 1	0.272	0.728	-8.581***
	57	Regression 2	0.363	0.637	-4.263***
Mechanical	213	Regression 1	0.298	0.702	-12.701***
	213	Regression 2	0.637	0.353	-3.682***
Transport	21	Regression 1	0.727	0.273	-1.343
	21	Regression 2	0.618	0.382	-1.386*

Statistically significant at the 1% level (***), at the 5% level (**), and at the 10% level (*)

Concerning the older and more mature engineering based sectors (Mechanical and Transport), the 'regression effect' has decreased significantly concerning Mechanical technologies, while this effect for Transport has been very low throughout the century and was not even statistically significant in the shift up to the war/postwar period. Concerning the 'mobility effect', it has also not been strong at all for the engineering based sectors, although it did manage to slightly exceed the convergence effect of the regression in recent times. These findings seem to indicate that for the older and more mature engineering based sectors less intra-group movements and integration of technological fields, as well as very little disruption in the evolution and compositions of technological opportunities, is to be expected in general.

Table 6.7. The extent of continuity (creative destruction or accumulative incrementalness) in the composition of technological opportunities within broad technological groups over time

Technological broad group	*Numbers of sectors*	*Period*	*Mobility Effect*		
			$\hat{\rho}$	$(1-\hat{\rho})$	$t_{\beta 0}$
Chemicals	50	Regression 1	0.480	0.520	3.786***
	50	Regression 2	0.105	0.895	0.728
Electrical / electronics	57	Regression 1	0.397	0.603	3.204***
	57	Regression 2	0.312	0.688	2.431**
Mechanical	213	Regression 1	0.349	0.651	5.411***
	213	Regression 2	0.407	0.593	6.466***
Transport	21	Regression 1	0.634	0.366	3.572***
	21	Regression 2	0.457	0.543	2.241**

Statistically significant at the 1% level (***), at the 5% level (**) and at the 10% level (*)

These overall results suggest that the evolution of twentieth century opportunities indicate patterns of uniform movements across different periods of technological development, rather than being marked by major disruptions (except for the science based sectors in recent times). Hence the composition of the technological opportunities, or sectoral growth rate ranking positions, does not tend to fluctuate across different historical waves of technological development. Together with a significant 'regression effect', this reflects that knowledge embodied in both older and newer fields of technological opportunities contribute to the development of a new technological paradigm. This evidence supports the now common view of creative incremental accumulation being dominant in recent times rather than radical shifts of 'creative destruction', and that new paradigms generally do not destroy old ones but complement and extend them (Pavitt, 1986; Patel and Pavitt, 1994).

Overall, this evidence supports the now common view of 'creative, incremental accumulation' with interconnected technological evolution. In this framework technologies previously following isolated channels of development are brought together, in a fashion in which knowledge from old systems is generally not destroyed in newer systems, but complemented and extended. This is what is causing blurring of technological opportunities in recent times. This is consistent with the evidence found in Pavitt (1986) and Andersen (2001, chapters 3, 6).

This 'creative incremental accumulation' in technological development does not, however, exclude other variables (such as institutions, corporate structures, the systems themselves, etc.) from experiencing creative destruction across technological epochs. (See Andersen 2001, chapter 6, on industrial dynamics, in which the co-evolution of technology and industry structures are examined. See also Andersen 2001, chapters 7 and 8, which documents how the structure of the technological profiles of firms becomes eroded across technological regimes).

5. REVEALED TECHNOLOGICAL TRAJECTORIES

Whereas the analysis so far has been based upon *overall* structures and the nature in which such structures change, this section aims to identify the individual trajectories of technological opportunities. The scope of this section is that, given the Schumpeterian institutional approach taken in this chapter, it is believed that the broad technological groups do not follow any general trajectory of changing technological opportunities (i.e. is gathered in only one trajectory type). However, it is expected that some trajectories are more likely to be followed rather than others owing to the 'instituted' nature of technological development, as defined by the existing technological paradigm and structure of socio-economic competence (Nelson and Winter, 1977; Dosi, 1982; 1988), see section 1. So rather than identifying randomly and unstructured the evolution paths of the selected 369 technological sectors (or patent classes), it seems more appropriate to examine to what extent a technological trend is typical.

Thus the next step of the analysis is to examine and identify how the underlying patterns of trajectories of technological opportunities, with respect to the individual 369 technological sectors, have evolved over time. This is carried out by deriving typical technological trajectories and selecting those of greatest historical technological importance.

5.1 The Model

Different possible trajectory types for the growth performance of patenting in each class can be identified using a general framework which has been developed for this purpose (see Table 6.8 developed from Table 6.1). In order to measure the evolution of trajectories, all Chemical, Electrical/electronics, Mechanical, Transport, and Non-industrial technological opportunities (or growth rates) have been ranked within a framework which sorts all 369 eligible technological sectors or patent classes into nine groups, derived by ranking them into three bands in accordance to their growth rates within each of three broad historical periods (interwar 1920–40, war/early postwar 1940–60, and recent period 1960–90). The growth rate ranking is carried out across all the broad technological groups, as in Table 6.1.

This scheme can then be used as a framework to reveal many different trajectory types showing alternative paths or ways in which the 369 eligible individual technological sectors may have changed their relative opportunities over time. Seventeen possible alternative trajectory types can be traced, when sorted in accordance with their initial starting point (whether they start ranked as high, medium, or low growth rate), and if they evolve in a linear or non-linear fashion. As appears below, there are also different types of linear and non-linear evolution paths.

Table 6.8. Framework for identification of trajectories of technological opportunities

Relative growth rate ranking position	*Historical time trend*		
	Interwar (1920–1940)	War/early postwar (1940–1960)	Recent period (1960–1990)
High	1	1	1
Medium	2	2	2
Low	3	3	3

Patent classes or technological sectors which follow a linear horizontal trajectory [1 (111); 2 (222); 3 (333)] reflect those technologies whose technological opportunities remain constant historically. Also, there are quadratic-like horizontal types of trajectories. If they are first declining, but then recovering [6 (121, 131); 7 (232)], these seem to be technologies with dropping opportunities in the war/early postwar period. However if they are first growing but then falling back to their initial starting point [12 (212); 13 (323, 313)] this indicates technologies with indeed great opportunities in the war/early postwar period. Finally there are historical declining and historical growing trajectories. Declining trajectories can be linear 5 (123), non-linear convex-like [10 (122, 133, 132); 11 (233)] or non-linear concave-like [16 (112, 113); 17 (223, 213)], and they all indicate historically falling opportunities. However, technological sectors performing growing trajectories, which can also be linear 4 (321), non-linear convex-like [8 (221, 231); 9 (332, 331)] or non-linear concave-like [14 (211); 15 (311, 322, 312)] indicate an historical growth in new opportunities.

Such an analysis of identifying typical technological trajectories will be carried out at the level of broad technology groups (i.e. Chemicals, Electrical/Electronics. Mechanical, and Transport, as defined in section 2).

The reason for carrying out such an analysis at the broad group level is mainly to adjust for size (i.e. number of technological fields or classifications), so that related technological sectors within the larger groups such as Electrical/electronics and Mechanical do not bias the results concerning which trajectories have been typical and dominant. This would display only Electrical/electronics and Mechanical major historical regimes. Hence, this analysis takes into account that what is interesting is not only the overall weight or significance of a trajectory for a technological group (Chemical, Electrical/electronics, Mechanical and Transport), but also a trajectory's relative typicality for this technology group. This is also another way of measuring a group's relative contribution to specific paths of development.

5.2 The Estimations

A trajectory's typical nature for a technological group relative to other trajectory types and in comparison to those that characterise other groups, can then be measured by an index termed the Revealed Technological Trajectory (RTT), first developed in Andersen (1998). The RTT index is a relative measure, and it can be compared to the Revealed Technological Advantage index (RTA) as well as the Revealed Comparative Advantage Index (RCA) which is commonly used in trade theory.

The value of the RTT index measures a broad technological group's share of technological sectors (or patent classes) following a particular technological trajectory type, relative to the overall share of all patent classes following this particular trajectory type. Hence, denoting by F_{Tj} the number of technological fields or sectors which follow technological trend T for a particular broad technological group j, the RTT index for each trend type in the broad technological group in question can then be defined as in equation (v).

(v)

$$RTT_{Tj} = \frac{F_{Tj} \Big/ \sum_T F_{Tj}}{\sum_j F_{Tj} \Big/ \sum_T \sum_j F_{Tj}}$$

where $F_{Tj}/\sum_T F_{Tj} \geq 0.1$ (10%)

The index varies around unity in such a way that when RTT > 1 the trajectory type in question is relatively typical for the technological group in question; hence the group has a relatively massive contribution to this particular technological development path or trajectory. However, RTT < 1 indicates that the trajectory in question is relatively uncommon for the technological group in question.

The RTT index of each of the 17 technological trajectory types has been calculated for all broad technological groups. However, only RTT results of trajectories which each accounts for at least 10% of its technological group's patent classes will be considered, to make sure that only trajectories which are of great overall historical importance for the broad technological groups are revealed.

5.3 The Results

The typical technological trajectories of great overall importance for each broad technological group (i.e. cases in which RTT>1 and which have an overall technological broad group patent share of about at least 10%) are displayed in Table 6.9. Table 6.9 shows that Electrical/electronics and Mechanical all have five different technological trajectories which are of great importance for the groups and typical relative to evolution paths in other groups, whilst Transport has four. However, Chemical technologies seem to be gathered in only three typical technological trajectories of great importance.

Table 6.9. Typical technological trajectories of great importance

Technological broad groups	*Number of important trajectories where RTT > 1*	*Trajectory types (See text for Table 6.8)*
Chemicals	3	1, 6, 16
Electrical/electronics	5	1, 6, 14, 16, 17
Mechanical	5	2, 3, 9, 15, 17
Transport	4	3, 4, 8, 9

These findings certainly suggest that it would be an oversimplification to draw any generalised conclusions about typical and historically important technological trajectories at the aggregate group level, because the different technological sectors belonging to each of the technological broad groups show a quite complex intra-group pattern in their technological evolution, with alternative typical technological trajectories going in totally different directions across different periods of technological development. Hence, all broad groups contribute to various alternative historical technological important paths of development. The typical trajectories of great importance for each broad technological group, as presented in Table 6.9, will now be explained.

Whereas the patent names at the patent class (or technological sector) level will be integrated in the text, some broader 56 technological groupings that the patent classes can be grouped into (in accordance with their common technological features) will be presented within 'apostrophes'.

Concerning Chemicals, as many as 44 out of the group's 50 eligible classes (i.e. 88%) are gathered in three typical technological trajectories showing historically important paths of development. In this context technological sectors documenting continuous opportunities throughout (trajectory 1) include the overall set of classes belonging to 'Agriculture Chemicals'. Also the overall set of classes belonging to 'Photographic Chemistry', and most patent classes within 'Synthetic Resins and Fibres' (including rubber, plastics, and adhesives from the polymer industry), as well as most classes belonging to 'Pharmaceuticals and Biotechnology' are gathered in this trajectory of continuous opportunities throughout; and so is a great part of the patent classes belonging to 'Chemical Processes' and 'Other Organic Compounds'. Another typical technological trajectory within Chemicals indicates sectors which drop out of the high opportunity ranking position in the war/early postwar period but then recover in later times (trajectory 6). This trajectory is most typical for patent classes belonging to chemistry of 'Inorganic compounds' as well as textile chemicals including 'Bleaching and Dying'. In fact, the overall set of classes within those latter mentioned groups follow this trajectory type. Finally, the last typical trajectory within Chemicals indicates classes which have enjoyed great technological opportunities until 1960 but then drop out in recent times (trajectory 16). This group includes half of the classes within 'Distillation Processes' (the other half already dropped out of high opportunities in 1940), as well as most classes within 'Coal and Petroleum Products' which especially peaked in development in the period ranging from the interwar and up to and including the early postwar period. Also about half of the sectors within organic compounds are gathered in this trajectory type.

The Electrical/electronics group's five typical technological trajectories of great importance for its technological group include 36 out of the sector's total of 57 eligible patent classes (i.e. 63%) selected for this analysis. The technological sectors' contribution to the specific paths of development indicated in the five typical trajectory types are in most cases spread across most of the 56 technological groupings. However, by investigating the specific classes within each of the groupings in relation to the five typical trajectories, and if any overall evolution trends have to be drawn, the five typical trajectories divide between sectors which have been of continuous importance throughout the century (trajectory 1), such as classes within 'Telecommunications', 'Office Equipment and Data Processing Systems' as well as 'Semiconductors', and other sectors which were just enjoying higher or greater opportunities at some technological epoch(s) of this century. Sectors with high opportunities in the beginning of the century, or growing in opportunities up to recent period but have since fallen in importance (trajectories 16 and 17), are typical sectors within the electrical equipment industry and technologies of ‘Electrical Devices and Systems’. Sectors which have risen in importance (such as trajectory 14) are typically sectors within ‘Communication Systems’ as well as ‘Other Electronic Devices’ including optics, LASER and space technology. Finally, sectors which show typical trajectories which indicate a drop in opportunities during the war/early postwar period (trajectory 6) cannot be generalised but spread across a broad range of different technological fields. However, about half of the classes within ‘Image and Sounds Equipment’ as well as ‘Photographic Equipment’ seem to follow this trajectory.

Within Mechanical only 109 out of the eligible 213 sectors in total (i.e. 51%) contribute to development path which are typical as well as important for the broad group. Hence this shows that we are dealing with a very technologically heterogeneous sector, which is hard to group into any particular or typical paths of development. Furthermore, within Mechanical only a few of the typical technological trajectories, and only some of the technological sectors within them, reflect technologies which have shown high technological opportunities at any point of time in the twentieth century; although Mechanical technologies as a group as a whole have been of great absolute importance in terms of accumulated technological size (or patent stock, see Graphs 1 to 4 in Figure 6.1). Also here, as within Electrical/electronics, the technological sectors’ contribution to the specific paths of development indicated in five typical trajectory types are not clustered in certain broader technological categories but spread across a range of the 56 technological groupings. Sectors belonging to those few trajectories which have shown increasing technological opportunities (trajectories 9 and 15) are, e.g., technologies belonging to ‘Miscellaneous Metal Products’, ‘Material Handling Equipment’, ‘Agriculture Equipment’, ‘Food, Drink and Tobacco Equipment’, ‘Other General Industrial Equipment’, ‘Power Plants’ etc. However, sectors which started with great opportunities but then indicated decreasing opportunities (trajectory 17) are e.g. ‘Other Specialised Machinery’ (wrapping, brushing, coating etc.), ‘Metal Working Equipment’, ‘Stone Working’, ‘Paper Making Apparatus’ etc.

The Transport group’s four typical and overall important technological trajectories presented in Table 6.9 include as many as 17 out of the sector’s 21 eligible patent classes (81%). A high proportion of Transport technologies seems to have been gathered in a trajectory with sectors continuously lowly ranked throughout this century (trajectory 3). These are particular technological sectors belonging to 'Railway and Railway Equipment' or technologies concerning wheels and axles within 'Transport Equipment'. However, other typical and overall important trajectories within Transport seem to show a rising tendency and an increase in technological opportunities (trajectories 4, 8 and 9) rather than a fall. The technological fields belonging to such growing trajectories are the overall set of classes within 'Internal Combustion Engines', the full set of classes belonging to 'Motor Vehicles', as well as a great part of the technologies within 'Ships and Marine Propulsion'.

From here it can be concluded that typical technological trajectories governing different technological sectors within each broad technological group (i.e., Chemicals, Electrical/Electronics, Mechanical, and Transport) explain technological evolution better than aggregate technological trajectories of broad groups as a whole, because all broad groups contribute to various alternative paths of development as captured by the RTT index.

It is here believed that sectors within each broad technological group that contribute to a specific path or trajectory of development which is typical and important for that group in question, may be related to 'families' of interrelated (i.e., complementary or co-evolving) technologies. Moreover, trajectories of technological opportunities that indicate similar development paths, or show similar opportunities across broad groups, may even be related to broader interrelated technological families. Yet, whether technological sectors grouped together within each of the typical technological trajectories, and whether typical trajectories within similar development paths across broad groups indicate interrelated technologies which could be interpreted as technological families require a much more elaborate analysis which is outside the aim of this Chapter.

However, based on another study (Andersen 2001, chapters 5 and 6), describing a century of technological opportunities in which the development paths of related technologies were investigated qualitatively, the results concerning such a relationship are very promising.

An example from that study will now be given in relation to the typical trajectories identified above. This example also documents how, after more isolated channels of development, the technological source sectors and diffusion sectors have become less focused and more complex over time.

That chemical engineering within the Chemicals broad technological group in the interwar period went through an epochal shift from coal based to petroleum–based feedstocks pushed a whole oil–based type of a paradigm up to and including the early postwar period, based on coal and petroleum products, distillation processes, and the development of new and better fuels for engines (trajectory 16); as well as a new range of materials from polymers (e.g., synthetic rubber, plastics, adhesives, manmade fibres (e.g.

nylon), Teflon, and many more) and other organic compounds whose opportunities have continued up to the present (trajectory 1). Similarly, the electrification and the development of electrical devices (trajectories 16 and 17) within the Electrical/electronics broad group up to and including the Early Postwar period have probably also been one of the most consequential technological changes. However, the more recent development of a new kind of paradigm of complex electronic based technologies (such as, e.g., electronic devices and related instruments including optics) (trajectory 14), as well as the continuous opportunities in technologies related to information and communication) (trajectory 1), would simply not have been possible without inventions and innovations within organic chemistry including the synthetic polymer industry. Freeman (1963) and Day (1990) argue how the polymer industry made electrical and electronic engineering manageable, and how it had essential applications in developing good electric insulators and advancing electrical and electronic engineering. Likewise, Chandler (1990, pp.217–221) emphasises how research in large companies, such as General Electric, became in direct competition with companies from the polymer industry, such as Du Pont, from their research on insulation for wire and moulding of carbon light bulbs, etc. Other studies on polymide applications in electronics include Grupp and Schmoch (1992) and van Vianen and van Raan (1992) who studied the crossroads in polymide chemistry and electronics focusing especially on LASER technology applied in medicine. This is just one example out of many concerning how technological families may have evolved and how technologies have become more complex. The evolution of possible 'waves' of interrelated innovation systems of generic technologies, and its co-evolution with industry structures, is examined wholly in Andersen (2001, Chapters 5 and 6), which deals with technological systems and industry dynamics.

Thus it can be argued that the quantitative results presented here, concerning typical technological trajectories of areas of greatest technological opportunities for each broad technological group, are not a purely statistical or random phenomenon, but match up quite nicely with what has been suggested in the literature on the history of technology and other case studies on technologies chosen for their particularly important contributions to development.

6. CONCLUSION

It can be concluded that the areas of greatest technological opportunities during the century ranging from 1890 to 1990 are not strongly concentrated within relatively few areas of related technological fields but have been increasingly widely dispersed across broad technological groups over different epochs of technological development; and that the evolution of the last century's technology can be divided into two major technological regimes or paradigms.

Whereas the first technological regime extending from the opening of this century until the war/early postwar period was characterised by intra-group technological diversification and the formation of a structure of specialised engineering and science–based fields, the one that has followed through to recent time (in which the gap between the science–based (Chemicals and Electrical/electronics) technologies on the one hand, and the engineering–based (Mechanical and Transport) technologies on the other, has been widening less quickly), is suggestive of an historical shift towards more integrated technological systems through the fusion of diverse and formerly separate branches of technology. The new paradigm governing the evolution paths of trajectories of technological opportunities builds to a greater extent on inter-group complementary and interrelatedness rather than on more isolated individual channels of development.

When exploring the nature by which technological evolution has increasingly been converted or channelled into wider-ranging and more complex technological systems, we saw how new systems are offshoots of a creative incremental technological development process in a variety of areas (as opposed to creative destruction) in which knowledge embodied in old technological fields within old systems is integrated in newer systems. This is witnessed by knowledge embodied in old paradigms being generally not destroyed but complemented and extended in new ones.

Finally, evidence shows how the Chemical, Electrical/electronics, Mechanical, and Transport broad technological groups' relative contribution to specific technological paths or trajectories have contributed to several alternative directions of opportunity development. This suggests that it is inappropriate to draw any 'general' conclusions concerning the specific contribution to the technological development of selected aggregate broad technological groups. This also demonstrates how revealed technological trajectories explain technological evolution better than the conventional aggregate measures that give an illusory picture.

Thus this chapter certainly supports the view that technology changes and trajectories evolve in an incremental, accumulative, and path-dependent fashion, and that some trajectories are more likely to be followed than others.

REFERENCES

Andersen, B. (1998). The evolution of technological trajectories 1890–1990. *Structural Change and Economic Dynamics*, 9 (1), 5–35.

Andersen, B. (2001). *Technological change and the evolution of corporate innovation. The Structure of Patenting*. Cheltenham: Edward Elgar.

Cantwell, J.A., Andersen, B. (1996). A statistical analysis of corporate technological leadership historically. *Economics of Innovation and New Technology*, 4 (3), 211–234.

Chandler Jr., A.D. (1990). *Scale and scope: the dynamics of industrial capitalism*. Cambridge, Mass.: Harvard University Press.

Day, L. (1990). *The chemical and allied industries*. In I. McNeil (Ed.), An Encyclopaedia of the history of technology. London: Routledge, 186–227.

Dosi, G. (1982). Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change, *Research Policy*, 11, (3), 47–162.

Dosi, G. (1988). Sources, procedures and microeconomic effects of innovation. *Journal of Economic Literature*, 26, 1120–1171.

Freeman, C. (1963). The plastics industry: a comparative study of research and innovation, *National Institute Economic Review,* 26, 22–61.

Freeman, C., Clark, J., Soete, L. (1982). *New technological systems: an alternative approach the clustering of innovation and the growth of industries*. In Unemployment and Technical Innovation (pp. 64-81). London: Frances Pinter Publishers.

Freeman, C., Perez, C. (1988). *Structural crises of adjustment: business cycles and investment behaviour.* In G. Dosi, C. Freeman, R. Nelson, G. Silverberg and L. Soete (Eds.), Technical Change and Economic Theory (pp. 38-66). London: Pinter Publishers.

Grupp, H., Schmoch, U. (1992). *At the crossroads in laser medicine and polymide chemistry: patent assessment of the expansion of knowledge*. In Hariolf Grupp (Ed.), Dynamics of science–based innovation (pp. 269-301). Berlin Heidelberg, New York: Springer Verlag.

Hart, P. (1994). *Galtonian Regression Across Countries and the Convergence or Productivity*. Discussion papers in quantitative economics and computing, University of Reading.

Hodgson, G. (1993). *Economics and evolution. Bringing life back into economics*. Cambridge: Polity Press.

Kodama, F. (1992). Technological fusion and the new R&D, *Harvard Business Review*, 92407, 70–78.

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: Chicago University Press.

Kuznets, S.S. (1930). *Secular movements in production and prices. Their nature and their bearing upon cyclical fluctuations* (Boston: Houghton Mifflin Co., The Riverside Press). 1967 reprint, New York: Augustus M. Kelley Publishers.

Nelson, R.R., Winter, S. (1977). In Search for a useful theory of innovation. *Research Policy*, 6, 36–76.

Nelson, R.R., Winter, S. (1982). *An evolutionary theory of economic change.* Cambridge, Mass: The Belknap Press of Harvard University.

Patel, P., Pavitt, K. (1994). The continuing, widespread (and neglected) importance of improvements in mechanical technologies. *Research Policy*, 23, 533–545.

Pavitt, K. (1986). *Chips and trajectories: how does the semiconductor influence the sources and direction of technological change?* In Roy M. MacLeod (Ed.), Technology and the human prospect. Essays in honour of Christopher Freeman (pp. 31-54). London and Wolfeboro: Frances Pinter.

Perez, C. (1983). Structural change and the assimilation of new technologies in the economic and social system. *Futures*, 15 (5), 357–375.

Rosenberg, N. (1976). *Perspectives on technology*. Cambridge: Cambridge University Press.

Schmookler, J. (1966). *Invention and economic growth*. Cambridge, Mass.: Harvard University Press.

van Vianen, B., van Raan, A.F.J. (1992). *Knowledge expansion in applied science: A bibliometric study of laser medicine and polymide chemistry*. In Hariolf Grupp (Ed.), Dynamics of science–based innovation (pp. 227–267).

von Tunzelmann, N. (1995). *Technology and industrial progress. The foundations of economic growth.* Aldershot: Edward Elgar.

Chapter 7

SCIENCE ON THE PERIPHERY: BRIDGING THE INFORMATION DIVIDE

Subbiah Arunachalam

M S Swaminathan Research Foundation, Chennai, India
E-mail: arun@mssrf.res.in

Abstract: Scientists in developing countries have access to only a tiny fraction of the information they need, and their own contribution to science is hardly noticed by others. It is important that these countries strengthen their scientific research and that their scientists become fully integrated members of the worldwide network of science. But, unfortunately, the transformations effected in the conduct of science with the advent of the new information and communication technologies (such as high bandwidth Internet) and the rising cost of subscriptions to journals and secondary services are widening the gulf between the industrialized and developing countries. Ironically, the steep rise in the cost of S&T information has helped developing countries in a way, as it forced scientists and librarians in the advanced countries to think of measures to overcome the 'serials crisis' and many of these measures can benefit developing country scientists. This paper looks at doing research in the developing world and how we can harness the new technologies to achieve information equity.

1. INTRODUCTION

Knowledge is the single most important commodity which drives socio-economic development in today's globalising world. Information and communication technologies (ICTs) are playing increasingly important roles in the production, transmission and utilization of knowledge. It was to understand the dynamics of ICT–mediated knowledge activities and to benefit from them that the OECD established the conference series called the Global Research Village (GRV). So far four conferences have taken place: Denmark, 1996; Portugal, 1998; The Netherlands, 2000, and Poland, 2002.

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 163-183.

Unfortunately what is referred to as global covers only the rich man's world – the OECD countries – and leaves out developing and least developed countries. GRVs are not the only conferences to deal with the issues of science at the global level. There was the Millennium Science Summit at Budapest, 2000, which gave pride of place to science in the developing world. But I doubt if the rhetoric will ever be translated into action. As pointed out by Bruce Alberts (1999), "most of the international organizations established by the United Nations with the great hope of using science and technology to improve the human condition are seriously hampered by bureaucracy and lack of energy, innovation, and resources."

Modern science and technology which have made the industrialized countries what they are today did not have the same transforming effect on the rest of the world. Benefits of modernity have not yet reached a very high proportion of populations in developing countries which suffer from persistent social, environmental and health problems. As Abdus Salam (1988), the Nobel Laureate in physics in 1979, observes, "in the final analysis it is basically mastery and utilisation of modern science and technology that distinguishes the South from the North". Part of the problem is caused by the pervasive privatisation of public knowledge in the advanced countries (Arunachalam 1996, Dickson 2003) — such as commercial publishers charging heavily for knowledge produced through public funding and pharmaceutical companies holding key healthcare knowledge wrapped up in patents. Is it at all possible for the developing countries, which have for long remained on the periphery, to become integrated into the 'Global Research Village'? Let us not have any illusions. Full integration is at best a long way off, if ever it can happen.

That does not mean we should abandon all hopes and refrain from doing something about it. This paper looks at doing research in the developing world and how we can harness the new technologies to achieve information equity. It begins with delineating the dimensions of the divide, as it affects the performance and utilization of scientific research, and goes on to describe efforts already afoot to bridge the divide, and finally lists possible courses of action. In particular, the paper discusses the implications of the new ICTs for research in the developing countries: Left to themselves, these technologies will only exacerbate the existing divides and make things worse. But the new ICTs also have the potential to bridge the divide and integrate science done everywhere. They can empower each individual scientist — for the first time — with the means to close the gap between the knowledge resources available in industrialized and developing nations.

2. SCIENCE ON THE PERIPHERY

Ideally, science is a truly global endeavour that knows no frontiers. Together with technology, science has long been a key driving force in the development process. In principle, anyone anywhere can contribute to the growth of knowledge in the sciences and make use of the collective knowledge, provided one has the inclination and capacity to do so. In the real world, production and efficient utilization of scientific knowledge are concentrated in a few industrialized countries. A large majority of countries — those on the periphery — contribute very little to the world's growing pool of scientific knowledge. In the years 1994–96, there was an average of 300 scientists and engineers (full time equivalent) per million of population in the South, as against the industrialized country average of 3,300 (UNDP 1997). There is also a great disparity in the investments made in science. General Motors invests far more in research than the entire R&D budget of India. The publication output of many developing countries is less than that of a single university department in advanced countries. In 1997, the developed countries accounted for some 84 % of the global investment in scientific research and development, had approximately 72 % of the world researchers, and produced approximately 88 % of all scientific and technical publications indexed in *SCI* (*Science Citation Index*) (UNESCO 2001). In essence, science on the periphery is characterised by poor funding, the absence of a viable scientific community, negligible presence in international invisible colleges, an insularity resulting from inadequate access to relevant information and inadequate communication within the local scientific community and with international invisible colleges, an unduly long time lag before participants in peripheral societies can take part in hot/emerging research fronts, lack of originality, weak institutional infrastructures, an excessive dependence on science carried out at the centre, and negligible contribution to the world's pool of knowledge. More importantly, it is rarely that scientists on the periphery take part in the collective endeavour of setting the research agenda in any discipline or research front (Arunachalam 1992).

Research effort is generally insufficient in most developing countries to have any real impact, and remains confined to a few sectors and a few disciplines. "The two main weaknesses exhibited by these countries — a lag in economic and social sciences, and weakness in engineering — make them continuously dependent on external assistance", points out the European Commission (1997). Unfortunately, at a time when international assistance is crucial, trends in official development assistance have been disappointing. Such "assistance by major donors in terms of their GDP declined from 0.33 % in 1990 to 0.22 % in 2000" (European Commission, 2003).

Not all developing countries are in the same boat. Some, such as India and China, have a large and reasonably developed S&T base. Brazil and South Korea are also catching up fast (Arunachalam 2002a).

3. SOME INDICATORS

Let us look at some literature–based evidence to realize the dimensions of the centre–periphery dichotomy in science. As seen from the *Web of Science*, only 15 countries had published more than 15,000 papers in the year 2002. The United States of America continues to remain the world's leading performer in science (with more than 240,000 papers indexed in 2002), followed by Japan, the United Kingdom, Germany, France, China, Italy, Canada, Russia, Spain, Australia, and India. Only two developing countries, the People's Republic of China (39,335 papers) and India (19,542 papers), appear in this list. In the Middle East Israel is the only country performing a substantial volume of scientific research (more than 10,500 papers). South Korea, Taiwan, and Brazil are three non-OECD countries that have published large numbers of papers.

In terms of papers published per unit population, the gap between the developed and the developing countries will be even more dramatic. As Frame *et al*. (1977) pointed out, the distribution of mainstream science production is even more skewed than the distribution of wealth amongst nations, with over 80 % of the world's mainstream scientific literature being produced by the top ten countries. In spite of the rapid strides made by countries such as China, South Korea, and Brazil (Arunachalam 2002a) and Latin America (Holmgren and Schnitzer 2004), the situation has hardly changed in most developing countries in the past quarter century. Some people have reservations in using *SCI* data for measuring publication output of developing countries.

Table 7.1. Percent share of world research in different fields for China and India as seen from different international databases

	TB [PubMed 1990– 1999]	CVD	Diabetes	New biol [BBCI]	Mathematics [Mathsci 2000]	Chemistry [CA 2001]	All science [SCI 2000]
India	5.34	0.66	1.11	1.35	2.02	2.5	1.55
China	1.11	1.04	0.63	2.03	10.35	9.8	2.83

Source: Arunachalam & Gunasekaran, 2002, Current Science*, 82, 933–947.*
TB=Tuberculosis. CVD=Cardiovascular diseases. CA= Chemical Abstracts. BBCI= Biochemistry & Biophysics Citation Index, 2000.

One can look at the numbers of papers indexed in subject–specific databases, such as *Mathsci*, *PubMed*, and *Chemical Abstracts*. Data for India and China are given in Table 7.1.

China publishes about 10 percent of the world's papers in both chemistry and mathematics. About ten years ago, China performed very much less science. The low volume of medical research carried out in both China and India is striking. It leads to the question of whether developing countries do research relevant to their needs. India accounts for more than 23 percent of the world's incidence of tuberculosis and China 17 percent, and yet their share of world research is very low (Arunachalam and Gunasekaran 2002a). India and China together account for over 26 percent of the prevalence of diabetes in the world and yet together they account for less than 2 percent of world's research in this area (Arunachalam and Gunasekaran 2002b).

As far back as 1990 cardiovascular diseases caused 2.3 million deaths in India and 2.6 million deaths in China in a year. Yet these countries together account for less than 2% of the world's research into cardiovascular diseases (Arunachalam & Gunasekaran 2001).

Table 7.2. Contribution to the world literature of tuberculosis of different countries and percent share of incidence of TB

Geographical entity	*No. of papers**	*% World share in research [A]*	*% World share in tuberculosis [B]+*	*Ratio A/B*
World	9,796			
USA	3,194	32.6	0.19	171.6
UK	1.311	13.4	0.08	167.3
G7	6,107	62.3	1.14	54.7
EU-15	3,563	36.4	1.57	23.2
Nordic countries	284	2.9	0.03	96.7
Australia	175	1.8	0.02	89.5
Israel	50	0.5	0.01	51.0
India	565	5.8	21.68	0.27
China	50	0.5	16.09	0.03
Brazil	116	1.2	1.40	0.84
Mexico	85	0.9	0.44	1.98
South Africa	393	4.0	2.46	1.63
Kenya	40	0.4	1.45	0.28

Source: Arunachalam & Gunasekaran, 2002, Current Science, 82, 933–947.
* Data from Science Citation Index 1998 [CD-ROM version].
+ Calculated from the data for the year 1999 provided by WHO, Global Tuberculosis Control, WHO Report 2001.

Whilst it will be unfair to expect China and India to perform medical research at the same level as the United States, it is reasonable to expect them to perform medical research at the same level as they perform mathematics and chemistry.

We compare the contributions of India and China to the world literature on tuberculosis and percent share of incidence of tuberculosis in these countries with those of some selected advanced countries in Table 7.2.

We used a set of carefully chosen keywords (tubercle, tubercul*, scroful*, Pott's and Mantoux) and made a basic index search of *SCI* (titles, keywords and abstracts) to identify papers on tuberculosis. The tremendous gap between what is needed and what is actually done by way of research in India and China is obvious from these figures. The ratio [% world share in research] / [% world share of disease burden] is even lower for India and China in the case of diabetes (Arunachalam & Gunasekaran 2002b).

Whilst much improvement can be brought about in the health sector in developing countries through better hygiene and sanitation and through application of knowledge already available, there is a clear need for research in areas where the advanced countries have no need or incentive to do research (except perhaps in drug development for markets in developing countries). Say Court and Young (2003), "While there is an extensive literature on the research–policy links in OECD countries, there has been much less emphasis on research–policy links in developing countries". Very little demand–led research in developing countries is reported and few systematic evaluations have been made of such research, says Louk Box [personal communication, 4 March 2004].

It is not just the volume of research carried out in developing countries and the relevance of the research to the immediate needs of these countries that are low but the impact of research and its influence on science *per se* and its applications are also low. One rarely sees papers from developing

countries in influential journals such as *Nature*, *Science* and *Proceedings of the National Academy of Sciences, USA*. Perhaps research and development in such countries needed to be relatively more 'applied and user–oriented' than in countries which were more developed. Moreover, even when some scientists from developing countries perform the kind of research that is acceptable for publication in such high impact journals, the exorbitant page charges levied by some of them are a major obstruction. Most papers from developing countries are published in low impact journals. They are, with some exceptions, hardly ever read or referred to by other researchers. In other words, much of science performed in these countries lacks visibility.

Scientists in developing countries, figuratively speaking, often live in islands of their own. Only rarely do advanced country researchers quote them, whereas developing country scientists quote a very large number of papers of advanced country scientists. Developing country scientists draw upon the rest of the world's knowledge, but their own contributions make hardly any dent on science performed elsewhere. What is more, whatever little is quoted of their work is quoted within the same discipline. Unlike in science at the cutting edge, there is very little interdisciplinarity in developing country science.

One characteristic feature of science today is the rising level of international collaboration. One simple measure of such collaboration is the number of papers jointly authored by researchers from different countries. Internationally coauthored papers nearly tripled in volume worldwide between 1986 and 1999, and in 1999, 17 % of scientific articles had at least one international author (National Science Board 2002). Much of this collaboration takes place amongst the OECD countries, especially the G7 countries. In contrast, very few papers result from international collaboration involving African and Latin American countries (Arunachalam 2000).

4. WHY SCIENCE IN THE DEVELOPING COUNTRIES?

Why should we be concerned about this skewed distribution of scientific research amongst nations? As science is a collective enterprise, the production of knowledge will be hampered if research is restricted to certain parts of the world and if only a selected few take part in the activity. Jawaharlal Nehru, free India's first Prime Minister, firmly believed, as did Abdus Salam, that even poor countries should invest in science. It was through research on high yielding varieties of wheat and rice that Monkombu Swaminathan and his colleagues in the Indian agricultural research establishment saved India in the late 1960s from recurring famine and transformed India into a food surplus country.

There is a great need for strong research institutions in developing countries. Bruce Alberts (2002) believes that every nation must have involved and effective institutions, run by the nation's own scientists and engineers, and he wants the scientists of the world to exploit fully the new communication technologies to share information and other resources which strengthen world science. Even the poorest nations must have scientists who are deeply involved in education at all levels, so as to produce the human capital on which much of development depends. "If we are to make long-term progress on our goal of producing a safer, more just world for our grandchildren, scientific capacity building and cooperative research programs deserve to be at the center of each of our international assistance programs", says Alberts (2002). Enhancing local S&T capacity is essential because trends in the development and use of new technologies have left a growing gap between the 'have' and 'have not' nations. The world has entered a vicious cycle in which developing countries which lag in S&T capacity are falling further behind, as industrialized nations with financial resources and a trained scientific work force exploit new knowledge and technologies more quickly and intensively (Alberts 2004). Says Kofi Annan (2003), "This unbalanced distribution of scientific activity generates serious problems not only for the scientific community in the developing countries, but for development itself". The InterAcademy Council (2004) has set out new initiatives for strengthening national scientific capabilities, not just those in the economically developed world, but in all countries, and for promoting global cooperation.

It is not always recognised by the developed world that the research generated in the developing countries is critical for the resolution of global problems such as infectious diseases, emerging new diseases, public health, environmental problems, climate change, biodiversity, taxonomy, and so on, says Barbara Kirsop of the Electronic Publishing Trust [personal communication]. These problems urgently need the scientific input from the developing countries that are so important for the global processes (Holmgren and Schnitzer 2004). Commenting on the US Treasury Department's ruling preventing American journal editors from editing manuscripts from certain countries thought to be hostile, Fox (2004) points to scientific advances in Cuba "in cancer therapeutics, in pediatric vaccines, and in humanized monoclonals" that American researchers cannot access. He classifies the decision as "blind ignorance" and muses: "By comparison, the Luddites were enlightened activists". A simple search on the Bioline International eprints server <bioline.utsc.utoronto.ca>, which already holds publications from 14 different developing countries, clearly demonstrates the significance of this missing science.

As the 10/90 report of the Global Forum for Health Research (2000) states, it is necessary for developing countries to develop the research capacity necessary for dealing with their own health problems through evidence based decision making. Less than 10 % of the worldwide expenditure on health research and development is devoted to the major health problems of 90 % of the population. There is a fundamental mismatch, expressed as millions of lives lost each year, between human needs and scientific innovation. If India and China are the leading sufferers of tuberculosis and diabetes in the world, it is in their own interest that they do part of the research to combat these diseases.

Consider tuberculosis. By the early 1990s, while TB was responsible for 2.8 % of the entire burden of ill health in the world, research on TB, at about

US $33 million in 1993, accounted for less than 0. % of the world's expenditure on health research and development. Funding for health research expressed as expenditure per DALY (disability adjusted life years) is ridiculously low for TB, viz. $0.68 per DALY in 1990 compared to asthma ($13.22), and blindness ($10.09). Second, TB in India and China is different from TB in the advanced countries of the West. The need for developing new drugs against *Mycobacterium tuberculosis*, in view of the ever growing emergence of new strains resistant to currently available drugs, and the limited efficiency of BCG vaccine against TB in adults and non-pulmonary forms of TB in India are additional reasons why India should pursue research in TB vigorously (Arunachalam & Gunasekaran 2002a). In the case of diabetes also, as Andrew Hattersley (2002) points out, what works elsewhere may not work in India, and vice versa, because environmental and genetic factors can make a difference.

There is one more reason. Multinational pharmaceutical companies are exploiting both the plant wealth and the knowledge of indigenous medical practitioners, and clever individuals and institutions in the industrialised countries are trying to apply for patents on the use of certain natural products, which are well known for centuries in some traditional societies (Arunachalam 1996). If the traditional societies rich in biological resources are to take full advantage of their natural wealth and not lose the advantage of having those natural resources, they ought to strengthen their research capacity. A strong science base can be the best defence in such situations.

5. INFORMATION KEY TO RESEARCH

We have seen that developing countries perform very little research and there is a genuine need for them to do much more. Now we shall look at the role of information and communication in research and how research in developing countries suffers for want of access to information.

Research is at once an intensely personal and a social activity. The aggregation and advancement of knowledge take place by collective efforts of researchers around the world. In the production of knowledge scientists use what is already known, and draw upon the knowledge generated by others across space and time. "It must be considered a birthright of scientific communities in a developing nation that the country should have at least one complete central library containing most of the scientific and technological journals, and all scientific books. There must be free access to this scientific literature", said Salam <http://wwwusr.obspm.fr/admin/seminaire/ chalonge/ dedic.html>. But unfortunately scientists in developing countries neither have access to large libraries nor do they have free access to most of the literature.

Generation of knowledge is only one part of the research process; for knowledge to be useful, it should be shared with other researchers and communicated, in a suitable format, to different users/ stakeholders. Every scientist also would like his/ her work to be used by others. Information and communication are two very important aspects of research. Scientists in developing countries are handicapped in both these respects.

Even when both information and communication were entirely mediated by the printed word, there was a big gap between the rich and poor countries that increased with the passage of time. Most journals were published in the West, and many libraries in developing countries could subscribe at best only to a few hundreds. Scientists in developing countries suffered a great deal of relative disadvantage compared to their western counterparts.

If escalating costs of print journals has made life difficult for scientists in developing countries, the advent of electronic sources of information has made the situation even worse. It is in the nature of any new technology to exacerbate the existing divide between the rich and the poor. The tremendous changes that are taking place in the way new information is published, stored, disseminated and retrieved using the rapidly advancing information and communication technologies have exacerbated the relative deprivation suffered by researchers in the developing world. The new ICTs have not just made each operation faster, but have brought about a greater synergy between these operations in ways unthinkable in the print era. Online publishing is raising the bar for resource discovery and is bringing extraordinary navigational ease at the desktop. In many institutions in developed countries one can now seamlessly move from one journal article to the full text of another referred to in the first article and from a reference in a bibliographic database to the full text of an article, thanks to agreements among journal publishers and database producers. For most developing country scientists this is still science fiction. The relative disadvantage of developing country scientists is higher now than in the print only era.

Many electronic journals accept manuscripts electronically and get the papers reviewed electronically. Many developing country scientists, who do not have access to personal computers, email, and the Internet, can neither submit their papers to these journals nor read them, nor act as referees. Nor are they able to take part in Internet– and grid–facilitated international collaborative projects. They are ‘excluded’.

The Internet access gap between the rich and the poor areas of the world is not only large, but is also growing (National Science Board 2002). In 1997 Internet host penetration rates in North America were 267 times greater than rates in Africa; by October 2000 the gap had grown to a multiple of 540.

ICTs have exacerbated the existing inequalities in the world in such a short time. Thanks to men like Gandhi, Martin Luther King, Nelson Mandela, and Desmond Tutu we have abolished skin–colour–based apartheid, but are letting the new ICTs create information–access–based apartheid.

One does not have to lose hope. As shown by the Information Village Research project of the M S Swaminathan Research Foundation in India, if intelligently and innovatively used the very same information and communication technologies can become an ally in our efforts in bridging the divides (Arunachalam 2002b). Even in the world of science and research. Let us see how.

6. EFFORTS AFOOT

The past few years have witnessed several developments which could make access to information for scientists in the developing world much more affordable. These include initiatives promoted by scientists, libraries, publishers, academies, and societies. Essentially these are of two kinds: one trying to bring down the access barrier or reducing the cost burden of libraries and the other enabling toll–free open access to information for all scientists. The first addresses largely the 'serials crisis' problem of librarians, and the second tries to help scientists and their institutions gain maximum mileage out of their publications. There are two kinds of initiatives under the second kind: open access journals and interoperable open archives. For want of space only a few examples are discussed briefly.

SPARC is the best example of the first kind. Backed by more than 600 member libraries of the International Scholarly Communications Alliance serving over 11 million students and faculty with a budget of US$5 billion, it aims to create high–quality low–cost alternatives to expensive commercial titles. It emerged from the widespread perception that in scientific communication the researchers and the laboratories — where scientific communication originates — have been forgotten or sidelined and that the profit motive of commercial publishers had taken over. SPARC's avowed aim is to restore the researcher in research publishing. SPARC persuades editors and editorial board members of expensive commercial journals to come out and start new journals of high quality. SPARC journals have become popular within a few years of their first publication. For example, the ACS journal *Organic Letters* has already registered a higher impact factor than its main commercial rival *Tetrahedron Letters*. Other SPARC journals which are doing well include *Theory and Practice of Logic Programming* and *Evolutionary Ecology Research*.

HINARI (Health Internet) [http://www.healthinternetwork.org], a UN/WHO initiative, provides free online access to commercial medical journals — more than 2300 of them from 47 publishers — to health professionals, medical researchers, and academics in 69 licensed countries in the developing world. An additional 44 countries may get access to the journals at a very low price. A total of 1,043 institutions in 100 countries (of a total of 113 eligible countries) have registered for the program (Aronson 2004). AGORA, sponsored by FAO with a view to promoting research and education in agriculture and related fields in the poor countries of the world, provides free or low cost access to 400 scientific journals in agriculture and related biological, environmental and social sciences to public institutions in developing countries. Although these programmes sound good, in reality though they may not be as good as they sound. The publishers do not provide free access to countries, such as India, where they already have a reasonably large subscription base, although India's per capita GDP is less than half of the US $1,000 agreed upon by WHO, FAO, and the publishers as the upper limit for being eligible to get free access to the journals! China and Pakistan also do not benefit from HINARI. Indeed, Elsevier, the largest commercial publisher in the field of science, technology and medicine, has three consortia subscriptions for *Science Direct* in India — signed with the Council of Scientific and Industrial Research, the Department of Atomic Energy and INDEST (a consortium of higher educational institutions). Ironically, many American universities are cancelling subscriptions to large aggregations of journals.

Elsevier Science, Academic Press, the American Physical Society, the Optical Society of America and World Scientific Publishing Company have joined ICTP/TWAS eJournals Delivery Service and have agreed to provide scientists in the South electronic access to their physics and mathematics journals. Launched in autumn 2001, eJournals Delivery Service now has more than 300 subscribers from over 60 developing countries. Subscribers have access to more than 240 journals. There are several other publishers who provide free online access to developing country scientists either immediately on publication or after the lapse of a few months. In a statement released in March 2004 in Washington DC, 48 not–for–profit publishers, representing more than 600,000 scientists and clinicians and more than 380 journals, pledged their support for providing free online access to their journals to scientists working in many low–income nations.

The African Virtual University is a 'university without walls' which uses modern ICTs to give institutions in sub-Saharan Africa direct access to some quality learning resources. It provides students and professionals in 17 countries free email accounts and access to an online digital library with over 1,000 full text journals. GDN Journal Services <http://www.gdnet.org/

online_services/journals/gdn_journal_services/index.html> works with partners to provide researchers working in developing countries with free access to a range of journals through GDNet.

These initiatives facilitate access to developed country journals for scientists in developing countries. What about helping developing country scientists gain greater visibility? There are a few non-profit publishers/distributors of developing country journals and information. These include Bioline International [www.bioline.org.br], which hosts electronic versions of many developing country journals; PERI (Programme for the Enhancement of Research Information) [www.inasp.info/peri/index.html], promoted by the INASP, a programme for the support to information production, access and dissemination for research partners in developing and transitional countries; SciELO [www. scielo.org], which hosts more than 80 journals published in Latin American countries and Spain; African Journals Online [www. inasp.info/ajol/ index.html], which provides free online access to titles and abstracts of more than 60 African journals and full text on request.

The Electronic Publishing Trust for Development (EPT), established in 1996, facilitates open access to the world's scholarly literature and supports the electronic publication of reviewed bioscience journals from countries experiencing difficulties with traditional publication. The EPT provides awareness of the benefits of electronic publishing, transfers e-publishing technology through training and online resources, provides management and distribution support, and supports open access initiatives and make them known to developing country scientists and publishers.

7. OPEN ACCESS

By ‘open access’ to the literature, we mean its free availability on the Internet, permitting any user to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the Internet itself (see www.soros.org/openaccess/read.shtml). The only constraint on reproduction and distribution is to give authors control over the integrity of their work and the right to be properly acknowledged and cited.

BioMed Central [www.biomedcentral.com] and the Public Library of Science [www.plos.org] are good examples of organizations publishing open access journals. BioMed Central (BMC) publishes about 100 journals, provides free access to all papers and encourages new free journals. It charges a handling fee of $500/article from the authors or their institutions (except those from the developing world). BMC publishes *Open Access Now*, a four page pullout in alternate issues of *The Scientist*, a fortnightly magazine for life scientists. It carries news and interviews relating to the open access movement. All papers published in BMC journals are automatically archived in PubMed Central (PMC). The US National Library of Medicine is scanning the back issues of PMC journals that are not already available in electronic form, and the complete contents will soon be available free in PMC.

The British Library is working to create a digital archive of electronic science publications. According to Sally Morris of the Association of Learned and Professional Society Publishers, many UK publishers are turning to the British Library to act as a centralized digital archive of papers.

The Public Library of Science (PloS) is a non-profit organization of scientists committed to making the world's scientific and medical literature freely accessible to scientists and to the public around the world for the benefit of scientific progress, education, and the public good. More than 3,000 scientists around the world have signed an open letter urging publishers to allow the research reports that have appeared in their journals to be distributed freely by independent, online public libraries of science.

PLoS launched its first online journal (*PLoS Biology*) in October 2003. An article on brain–machine interface [DOI: 10.1371/journal.pbio.0000042] attracted thousands of hits and downloads within a few days leading to the crashing of the server! The incident shows how popular the idea of open access is. *PLoS Medicine* will be launched in 2004. The costs of peer review, editorial oversight, and publication are recovered from the authors.

Although *Journal of Clinical Investigation* has been open access since 1996 (Savla 2004) and has a high impact factor (14.051 for 2002), it was *PLoS Biology* that received much public attention.

Of the estimated 24,000 scientific journals, more than 820 are now (as of April 2004) open access [www.doaj.org] including many developing country journals. All 10 journals of the Indian Academy of Sciences, for example, are open access. *BMJ* was among the earliest to go open access. The National Academy of Sciences, USA, is a model for other societies and academies. Its entire collection of over 2,900 reports and books are available free on the Web to users in developing country, and its *Proceedings* can be accessed free from many developing countries. Indeed, NAS President Prof. Bruce Alberts is a champion of the cause of science and technology in developing countries. Alberts (1999) suggests the following two–part strategy:

– Connecting all scientists to the World Wide Web, where necessary by providing subsidized Internet access through commercial satellite networks; and
– Taking responsibility for generating a rich array of scientifically validated knowledge resources, made available free on the Web, in preparation for a time when universal Internet access for scientists is achieved in both developing and industrialized nations.

Both of these are excellent suggestions. Not only do we need useful content to be available free on the Web, but we also need the technology in place to take advantage of the content. We should persuade philanthropic foundations and donor agencies concerned with higher education and research to donate funds to make PCs and high bandwidth Internet connections available to researchers and libraries in developing countries. The concerned governments should make things easy for the spread of ICTs among university and research institutions.

8. OPEN ARCHIVES

Let us now turn our attention to initiatives on providing toll–free access to information through the interoperable archives. The full–text physics archive, arXiv, founded by Paul Ginsparg at Los Alamos in 1991 and now moved to Cornell University, is probably the oldest subject–specific eprint archive. This pioneering effort is easily one of the most innovative and successful experiments to date in scholarly communication. With more than fifteen mirror sites around the world including five in Asia, one in Brazil and one in South Africa, this automated electronic archive provides free access to research papers in physics, mathematics, nonlinear sciences, computer science, and quantitative biology. arXiv contains more than 267,000 papers (as on 8 March 2004), of which about a half are in astrophysics and high energy physics. Another physics database is the SPIRES HEP literature database, which has more than 500,000 high energy physics related articles, including journal papers, preprints, eprints, technical reports, conference papers and theses, comprehensively indexed by the SLAC and DESY libraries since 1974. The Astrophysical Data system, [http://adswww.harvard.edu], hosted by the Harvard–Smithsonian Center for Astrophysics and funded by NASA, maintains four bibliographic databases containing more than 3.6 million records.

ResearchIndex (or CiteSeer), the full–text archives for computer science, founded by Steve Lawrence of NEC Research, Princeton, NJ, is a scientific literature digital library that aims to improve the dissemination and feedback of scientific literature, and to provide improvements in functionality, usability, availability, cost, comprehensiveness, efficiency, and timeliness. It has more than 400,000 papers. It is more than a digital library; it provides algorithms, techniques, and software that can be used in other digital libraries. ResearchIndex indexes Postscript and PDF research articles on the Web. It autonomously creates a citation index that can be used for literature search and evaluation. ResearchIndex uses search engines and crawling to locate papers efficiently on the Web. Authors need not submit thair papers in any special format. The full source code of ResearchIndex is available at no cost for non-commercial use.

Cogprints, founded by Stevan Harnad at the University of Southampton, UK, is an electronic archive for all papers that are pertinent to the study of cognition. It runs on EPrints open archive software, a freely distributable archive system available from EPrints.org. The open archives interoperability is achieved using open archives protocol developed by the Open Archives Initiative (OAI) <http://www.openarchives.org> at Cornell university. Bioline International [http://www.bioline.org.br/] has launched the Bioline International EPrints Repository, an open–access, OAI-compliant archive for bioscience, especially from developing countries. Clinmed [clinmed. netprints.org], launched in December 1999 as a collaborative venture of the BMJ Publishing Group and Stanford University Libraries' HighWire Press, is a website that provides a place for authors to archive their completed original research into clinical medicine and health — before, during, or after peer review by other agencies. All articles fulfilling certain minimal conditions will be posted, usually within 24 hours of receipt. There are similar services in economics (RePEc) and computing (CoRR).

According to Kat Hagedorn of OAIster, there were more than 1.5 million papers in open archives by the end of 2003.

Electronic eprints do not aim merely at capturing the articles; it is far more than a simple electronic reproduction of what would appear in print journals. Eprint publication on the web offers numerous value–added elements, such as multimedia, data sets, as well as contextual links to other documents referred to in a paper and to databases. Indeed, the document linking advantage is being exploited by digital libraries, commercial aggregators of journals and secondary service providers such as Thomson–ISI and the *Chemical Abstracts* Service. In the very near future the print versions of journals will not be the true archivers. The eprint archives, as both the data and the access systems can be mirrored in several locations around the world, offer in-built insurance against possible loss of archived material due to unforeseen calamities (such as natural disasters or system failures at any one location).

9. WHAT NEEDS TO BE DONE?

The arguments are clear and simple that open access has many advantages and hardly any drawback, and yet many developing country scientists are not embracing it. The eloquent writings of champions such as Stevan Harnad and Peter Suber have not had the desired results so far. What could be the reason? Is it sheer lethargy or inertia to change the way people do things or is it ignorance about the advantages of open access (institutional eprint archives, open access toll–free journals)?

The results of a recent author survey funded by the Joint Information Systems Committee, UK, and the Open Society Institute show that a low percentage of authors have deposited pre-prints or post-prints in institutional repositories (IR), although a very high percentage would do so if required to do so by their funder or employer. Funding agencies may wonder whether IRs offer a sustainable high quality service for the record of science. The perception may be that IRs will not hold the best version of a research article, that they will not be indexed as comprehensively as traditional journals, and that their long-term preservation is not secure. To some researcher IRs appear to be 'anarchic', whereas journals — whether subscription or OA — are perceived to be organized, reliable and secure [Frederick Friend, personal communication]. All of them untrue. Recently Thomson–ISI announced that they will be indexing OA papers in the *Web of Science*, and as of April 2004, more than 190 OA journals were indexed in the *Web of Science*. In the UK the House of Commons Science and Technology Committee is looking into scientific publications and is gathering the views of all stakeholders. It is rather unfortunate that the focus of the public hearings is on 'publishing' and not on the larger and more relevant issue of 'access'.

There is some discussion on the need for donor agencies and governments to insist that papers resulting from funds provided by them should be made freely available to all either through placing them in interoperable archives or by publishing them in open access journals. Indeed Congressman Martin O. Sabo introduced a bill (The Public Access to Science Act) in the US House of Representatives on 26 June 2003 that would make research funded by the American government exempt from copyright protection.

Says Stevan Harnad [https://mx2.arl.org/Lists/SPARC-IR/Message/145-P.txt, 26 November 2003]: "If we are to have open access to all refereed research, researchers have to be persuaded (or obliged) to provide it. … The only thing that will persuade researchers to provide open access is a powerful and irrefutable empirical demonstration of the fact that doing so is in their own interests — indeed, a demonstration of 'how much' it is in their own interests, and how much they (and their institutions) are losing, daily, monthly, yearly, until they do provide open access to their refereed research output".

Harnad and colleagues from four universities (Quebec at Montreal, Southampton, Oldenburg and Loughborough) are working on producing the empirical demonstration — of the direct causal connection between research access and research impact, and the substantial size of the benefits. They are comparing articles in the same (toll access) journal that have and have not been made OA by self–archiving. Preliminary results for physics and computer science show that OA papers are cited more often than non-OA papers [http://www.ecs.soton.ac.uk/~harnad/Temp/OA-TAadvantage.pdf]. When the results are published they hope to persuade not only the researchers to provide open access, but their employers and funders to extend their existing 'publish–or–perish' policies to mandate that their researchers provide it. Estimates made of the average number of downloads of papers deposited with arxiv [http://opcit.eprints.org/ijh198/13.html] encourage open access.

Steve Lawrence (2001) has shown that articles (in computer science and related fields) freely available online are more highly cited and suggests that for greater impact and faster scientific progress, authors and publishers should aim to make research easy to access. Alonso and Fernandez–Juricic (2002) have shown that the impact factor of a set of Latin American journals have more than doubled after they became freely available online through SciELO.

The Open Society Institute [http://www.soros.org/openaccess], has issued a statement advocating open access and has provided $3 million over three years for projects supporting 'alternative' journals and open archiving initiatives. In June 2003, a group of researchers met at the Howard Hughes Medical Institute, and released the Bethesda Statement on Open Access Publishing [http://www.earlham.edu/~peters/fos/ bethesda.htm]. In October 2003 representatives of many European organizations met in Berlin and drafted the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. Francis Muguet, a champion of open access, and CERN organised two seminars on open access at the World Summit on the Information Society, Geneva, in December 2003.

Unfortunately much of the discussion on open access currently focuses on OA journal publishing and the 'author pays' model of sustaining these journals. We should shift the focus to the far more useful and simpler option of OA archives that would bring in millions of papers into the public domain at insignificant costs (Arunachalam 2004). Developing country scientists should immediately adopt the unified dual open access provision policy:

1. Publish your article in a suitable open–access journal whenever one exists [http://www.earlham.edu/~peters/fos/boaifaq.htm#journals].
2. Otherwise, publish your article in a suitable toll access journal and also self–archive it [http://www.eprints.org/self-faq].

If there is no institutional archive then authors may place their papers in existing archives.

10. CONCLUSION

What we need to achieve is to make scientific and technical information flow freely and be accessible at affordable costs to researchers and students everywhere in the world. A kind of enlightened socialism, as it were, for scientific knowledge. To be honest, this could only be an ideal — the direction in which we should move. Achieving this goal would necessarily mean exploring many possibilities. First, we should try to facilitate access to all the content (scientific and technical journal papers, reports, theses, conference papers, bibliographic, factual and full–text information, etc.); second, we should ensure that all potential users of this content have access to the technological tools for accessing it (such as computers and high bandwidth Internet connection); third, we should continue with our efforts to evolve standards and norms, including research into better ways of doing things, that will enhance the ease of use and value of the content; fourth we should build organizational structures that would ensure the long term survival of the entire system.

Unfortunately, there is a tendency for privatisation of knowledge and much resistance to letting knowledge pass on to public domain. Scientists and institutions in developing countries should forge alliances with forces which want to democratise knowledge and persuade their counterparts in the advanced countries and international bodies to support efforts that will make the playing field level.

ACKNOWLEDGEMENT

I am grateful to Barbara Kirsop of the Electronic Publishing Trust and Stevan Harnad of the University of Southampton for their valuable suggestions.

REFERENCES

Alberts, B. (1999). *Science and the world's future*. President's address to the Fellows of the National Academy of Sciences, USA, Washington D.C., 26 April 1999.

Alberts, B. (2002). *Engaging in a worldwide transformation: our responsibility as scientists for the provision of global public goods*. President's address to the Fellows of the National Academy of Sciences, USA, Washington D.C., 29 April 2002.

Alberts, B. (2004). The InterAcademy Council: Inventing a new global organization. *The Scientist*, 18 (4) (1 March).

Alonso, W.J., Fernandez–Juricic, E. (2002). Regional network raises profile of local journals. *Nature*, 415, 471–472.

Annan, K. (2003). A challenge to the world scientists*. Science*, 299, 1485.

Aronson, B. (2004). Improving online access to medical information for low–income countries. *New England Journal of Medicine*, 350, 966–968.

Arunachalam, S. (1992). *Periphery in science: What should be done to help peripheral science get assimilated into mainstream science?* In R. Arvainitis, J. Gaillard (Eds.), Science Indicators for Developing Countries. Paris: Orstom.

Arunachalam, S. (1996*). Science on the periphery enriches mainstream science but at what cost? The case of ethnobotany*. In R Waast (Ed.), Les Sciences Au Sud E'tat Des Lieux: Science in the south (Vol. 6, pp. 29–50). Paris: Orstom.

Arunachalam, S. (2000). *International collaboration in science: The case of India and China*. In B. Cronin B, H.B. Atkins (Eds.), The Web of Knowledge: A Festschrift in honor of Eugene Garfield (pp. 215–231). Medford: Information Today Inc.

Arunachalam, S. (2002a). Is science in India on the decline? *Current Science*, 83, 107–108.

Arunachalam, S. (2002b). Reaching the unreached: How can we use information and communication technologies to empower the rural poor in the developing world through enhanced access to relevant information*? Journal of Information Science*, 28, 513–522.

Arunachalam, S. (2004). India's march towards open access. [http://www.scidev.net/Opinions/index.cfm?fuseaction=readOpinions&itemid=243&language=1].

Arunachalam, S., Gunasekaran, S. (2001). *Cardiovascular diseases research in India and China in the 1990s*. In Proceedings of the 8th International Conference on Scientometrics and Informetrics (pp.53–62). Sydney: University of New South Wales.

Arunachalam, S., Gunasekaran, S. (2002a). Tuberculosis research in India and China: From bibliometrics to Research Policy, *Current Science*, 82, 933–947.

Arunachalam, S. Gunasekaran, S. (2002b). Diabetes research in India and China today: from literature–based mapping to health–care policy. *Current Science*, 82, 1086–1097.

Court, J., Young, J. (2003). *Bridging Research and Policy: Insights from 50 case studies*. ODI Working Paper 213. London: Overseas Development Institute. (available at www.odi.org.uk/publications).

Dickson, D. (2003). The threat to science as a public good. <http://www.scidev.net/Editorials/index.cfm?fuseaction=readEditorials&itemid=55&language=1>.

European Commission. (1997). EUR 17639–*Second European Report on S&T Indicators*. Luxembourg: European Commission.

European Commission. (2003). EUR 20025*–Third European Report on Science & Technology Indicators*. Luxembourg: European Commission.

Fox, C.A. (2004). Scientific censorship. *Chemical and Engineering News*, 82 (9), 4–5.

Frame, J.D. , Narin, F., Carpenter, M.P. (1977). World distribution of science. *Social Studies of Science*, 7, 501–516.

Global Forum for Health Research. (2000). *The 10/90 Report on Health Research*.

Hattersley, A.T. (2002). Multiple facets of diabetes in young people. *Current Science*, 82, 273–278.

Holmgren, M., Schnitzer, S.A. (2004). Science on the rise in developing countries. *PLoS Biology*, 2, 10–13.

InterAcademy Council. (2004). *Inventing a better future: a strategy for building world capacities in science and technology*. Amsterdam: IAC.

Lawrence, S. (2001). Online or invisible? *Nature*, 411, 521.

National Science Board. (2002). *Science and Engineering Indicators –2002*. Arlington, Va.: National Science Foundation.

Salam, A. (1988*). Notes on science, technology and science education in the development of the South* (Prepared for the 4th Meeting on the South Commission, 10–12 December 1988, Kuwait). Trieste: The Third World of Academy of Sciences.

Savla, U. (2004). Open access already exists*.* *Science*, 303 (5663), 1467.

UNESCO (2001). *The state of science and technology in the world, 1996–1997*. Montreal: UNESCO Institute for Statistics.

United Nations Development Programme, 1997, *Human Development Report 1997*. Oxford University Press, New York. See Human Development Indicators, Table 15: Educational Imbalances, p. 81.

Chapter 8

DATA MINING AND TEXT MINING FOR SCIENCE & TECHNOLOGY RESEARCH

Edda Leopold, Michael May, and Gerhard Paaß

Fraunhofer Institut Autonome Intelligente Systeme, Sankt Augustin, Germany
E-mail: edda.leopold@ais.fraunhofer.de

Abstract: The goal of the paper is to give an overview on the state of the art of data mining and text mining approaches which are useful for bibliometrics and patent databases. The paper explains the basics of data mining in a non-technical manner. Basic approaches from statistics and machine learning are introduced in order to clarify the groundwork of data mining methods. Text mining is introduced as a special case of data mining. Data and text mining applications especially useful for bibliometrics and querying of patent databases are reviewed and three case studies are described.

1. INTRODUCTION: DATA MINING, TEXT MINING, INFORMATION RETRIEVAL

'Knowledge discovery' or 'data mining' is the partially automated process of extracting patterns or models from usually large databases. As a scientific discipline it is a relative newcomer, building on work in many areas such as machine learning, statistics, information retrieval, and database technology. From those individual disciplines it can be distinguished by its aim of integrating the various individual lines of research and by its stronger emphasis on technology and applications (Hand et al., 2001). The latter makes it a highly relevant topic for science and technology research.

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 187-213.

In this paper we introduce basic approaches from statistics and machine learning in order to clarify the groundwork of data mining methods. Text mining which is especially relevant for science & technology research is introduced as a special case of data mining. The data mining techniques which are applied to text have to deal with the specific qualities of textual data: skewness of frequency distributions, synonymy, polysemy, sparseness of data, etc. This has important implications for the choice of data mining techniques which are suitable to the textual domain. Furthermore, there are techniques which are specific to text mining but are not covered by data mining proper (see section 3).

Information retrieval (IR) is a discipline which dates back to the seventies. It deals with the representation, storage, organisation, and access to information items (Baeza-Yates and Ribeiro-Neto, 1999). Given a user's query the goal of an information retrieval system is to retrieve information which might be useful or relevant to the user. Many techniques which have been developed in IR are nowadays employed in the area of text mining.

Text mining offers a variety of approaches for extracting information and knowledge from textual data:

- The *classification* of unlabelled text documents into a set of predefined classes can be used for the generation of ontologies and semantic spaces.
- *Clustering* of unlabelled documents according to their similarity can be deployed for the detection of related information.
- *Semantic spaces (also called topic maps)* can be utilised to obtain an insight into the semantic relations between documents in a document collection.
- The *segmentation* of texts into thematically coherent units and methods for detecting new and emerging topics in text documents allow a more efficient access to textual information.
- The *attribution of authorship* based on lexical and syntactical characteristics of the text can be used the for detection of plagiarism and therefore has implications to the management of intellectual property rights.
- *Named entity recognition* permits the search for persons, organisations, chemical substances and the like in textual data.

For all these tasks the internet and emerging topics such as the semantic web pose new challenges. After giving a survey on the various approaches (section 2), more concrete case studies are given to illustrate the data mining approach and its usefulness for bibliometrics and the management of patent databases (see section 4). It will be shown how web documents can be automatically inserted into a predefined ontology using text classification techniques. It will be illustrated how text mining can used to clarify disputed authorship and it will be described how semantic spaces can be used in order to refine the categorisation of an existing database of pre-classified documents.

2. DATA MINING FUNDAMENTALS

Data mining methods can be roughly divided into *supervised* and *unsupervised* methods. Supervised data mining methods learn from training data whereas unsupervised data mining methods use other cues such as, the Euclidean metric on input space.

In order to apply these methods to textual documents we have to represent them as numeric vectors, which can be readily processed by statistical estimation procedures. Surprisingly it is sufficient for many applications to simply count the number of occurrences of each word in a document, the so called *bag–of–words* representation (Salton and McGill, 1983). To indicate the utilisation of data mining methods for text mining, we will describe the different procedures using this representation. Later we will discuss alternative representations of documents in more detail.

The *supervised* data mining methods which we are going to present classify documents into predefined classes. This means they can be used to insert new documents into an already existing ontology. The *unsupervised* data mining methods cluster texts according to their (semantic) similarity or reduce the dimensionality of text representations.

2.1 Supervised Data Mining Methods

This section discusses advantages and disadvantages of the most important supervised learning methods used in data mining. During the last few years a number of estimation techniques have been proposed and evaluated which may be used for data mining and text classification. An example would be to automatically assign each incoming publication to a classification code of the *Science Citation Index* (SCI) such as ‘sport’, ‘politics’, or ‘arts’.

Whatever the specific method employed, a supervised data mining task starts with a *training data set*. A data mining method is given a set of instances (text documents) which are already labelled according to the class they belong to (‘sport’, ‘politics’, etc.). The task is then to learn a model based on the information provided in the training data (the words contained in the document) such that the document can be classified into one of the categories based on that information.

In the second step a *test data set* with the same general structure is given to the algorithm. However, the class information, although known to the user, is hidden from the algorithm. Using the model built in the first step, the algorithm classifies the instances into the predefined categories. Many more complex variants of these basic scheme exists, but the division into training and test data, where the model is built on the training set and its performance evaluated on the test set is common to most of them. This form of learning is called *supervised* learning, because the learning process is guided by the already known class information. The performance of the algorithm is measured by determining how successful the algorithm is in predicting the unknown class.

There are different assessment methods which measure the performance of a classifier. Suppose that there are two classes of documents in a document collection, a positive class ('sport') and a negative class ('non-sport'). Let tp and tn denote the number of documents that the classifier correctly identifies as positive and negative respectively, and let fp and fn denote the number of documents that are wrongly classified an positive or negative. *Precision* (prec.) and *recall* (rec.) are defined as follows:

$$\text{rec.} = \frac{tp}{tp + fp} \quad \text{prec.} = \frac{tp}{tp + fn}.$$

In terms of Information Retrieval recall indicates how many of the relevant documents are retrieved and precision quantifies how many of the retrieved documents are in fact relevant. Obviously there is a trade off between precision and recall. When an IR system searches restrictively it may retrieve a few irrelevant documents, therefore precision is high. However, many relevant documents might have been overlooked, which corresponds to a low recall. When, on the other hand, the search is more exhaustive, recall increases and precision goes down. The *F*-score is a compromise between recall and precision for measuring the overall performance of a supervised classifier. It is defined as

$$F_\alpha = \left(\frac{\alpha}{\text{prec.}} + \frac{1-\alpha}{\text{rec.}} \right)^{-1},$$

where α is a factor which determines the weighting of precision and recall. A value of $\alpha = 0.5$ is often chosen for equal weighting of precision and recall. Accuracy (acc.) and Error (err.) are further assessment methods. They measure the fraction of correctly (or wrongly) classified documents in relation to the total number of documents. More formally

$$\text{acc.} = \frac{tp+tn}{tp+fp+tn+fn}, \quad \text{err.} = \frac{fp+fn}{tp+fp+tn+fn}.$$

Error and Accuracy are inappropriate performance measures for most text mining tasks, because the number of documents in the negative class is usually very large and so is the number of correctly classified negative documents. Therefore tn is large, which makes Accuracy less sensitive to the small but interesting quantities tp, fp, and fn. (Manning & Schütze 1999)

2.1.1 Naïve Bayes

Probabilistic classifiers rely on the assumption that the words of a document d_i have been generated by a probabilistic mechanism. For classification only the influence of the underlying class such as 'sports' or 'politics' is of interest. Therefore it is assumed that the class $c(d_i)$ of a document determines the probability $p(w_1,\ldots,w_N|c(d_i))$ of its words. Now we may use the *Bayesian Formula* to determine the probability of some class if the words $w_1,\ldots,w_N$ of a document are known

$$p(c_m \mid w_1,\ldots,w_N) = \frac{p(w_1,\ldots,w_N \mid c_m)p(c_m)}{\sum_{k=1}^{K} p(w_1,\ldots,w_N \mid c_k)p(c_k)}.$$

Note that the documents may belong to one of K different classes. The *prior probability* $p(c_m)$ denotes the probability that some arbitrary document belongs to class c_m before its words are known. Often the prior probabilities of all classes may be taken to be equal. The conditional probability on the left is the desired *posterior probability* that the document with words $w_1,\ldots,w_N$ belongs to the class c_m. We should assign the class with the highest posterior probability to our document.

For document classification it turned out that the specific order of the words in a document is not very important. Even more, we may assume that for documents of a given class a word appears in the document irrespective of other words. This leads to a simple formula for the probabilities of words

$$p(w_1,\ldots,w_N \mid c_m) = \prod_{n=1}^{N} p(w_n \mid c_m).$$

Combining this with the Bayes formula defines the Naïve Bayes classifier. Simplifications of this sort are required because many thousand different words occur in a corpus.

The naïve Bayes classifier involves a training step which simply requires the estimation of the probabilities of words $p(w_n|c_m)$ in each class by its relative frequencies in a training sample. In the *classification step* the estimated probabilities are used to classify a new instance according to the Bayes rule. Although this model is unrealistic it yields surprisingly good classifications (Dumais et al. 1998, Joachims 1998). In contrast to other classification approaches it estimates the probabilities of classes. It may be extended in several directions (Lewis 1998; Sebastiani 2002).

2.1.2 k-Nearest Neighbour

Instead of building explicit models for the different classes we may select training documents which are "similar" to a test document. The class of the test document subsequently may be inferred from the class labels of the "similar" training documents. If k similar documents are considered the approach is also known as *k-nearest neighbour classification*. There are a large number of similarity measures used in text mining. If w_{in} is the count of the n-th word in a document d_i the *cosine similarity measure* of documents d_i and d_j is defined as

$$S(d_i, d_j) = \sum_{n=1}^{N} w_{in} w_{jn} \Bigg/ \sqrt{\sum_{n=1}^{N} w_{in}^2} \sqrt{\sum_{n=1}^{N} w_{jn}^2} .$$

Other similarity measures are discussed in (Baeza-Yates & Ribeiro-Neto 1999). For deciding whether a document d_i belongs to a class c_m the similarity $S(d_i,d_j)$ of all documents d_j in the training set is determined. The k most similar training documents (neighbours) are selected. The proportion of neighbors having the same class may be taken as an estimator for the probability of that class. If the largest proportion exceeds some threshold the corresponding class is assigned to the document d_i. The threshold as well as the optimal number k of neighbours may be estimated from additional training data by cross-validation.

Nearest neighbour classification is a non-parametric method and it can be shown that for large datasets the error rate of the 1-nearest neigbor classifier is never larger than twice the optimal error rate (Hastie et al., 2001). Several studies have shown that nearest neighbour methods have very good performance in practice (Joachims, 1998; Yang, 1999). Their drawback is the computational effort during classification, where basically the similarity of a document with respect to all other documents of a training set has to be determined. Some extensions are discussed in (Sebastiani, 1991).

2.1.3 Support Vector Machines

A Support Vector Machine (SVM) is a supervised classification algorithm which recently has been applied successfully to text classification tasks. SVMs have proved to be an efficient and accurate text classification technique (Joachims, 1998; Dumais et al., 1998; Drucker et al., 1999, Leopold and Kindermann, 2002). Like other supervised machine learning algorithms, an SVM works in two steps. In the first step — the *training* step — it learns a decision boundary in input space from preclassified training data. In the second step — the *classification* step — it classifies input vectors according to the previously learned decision boundary.

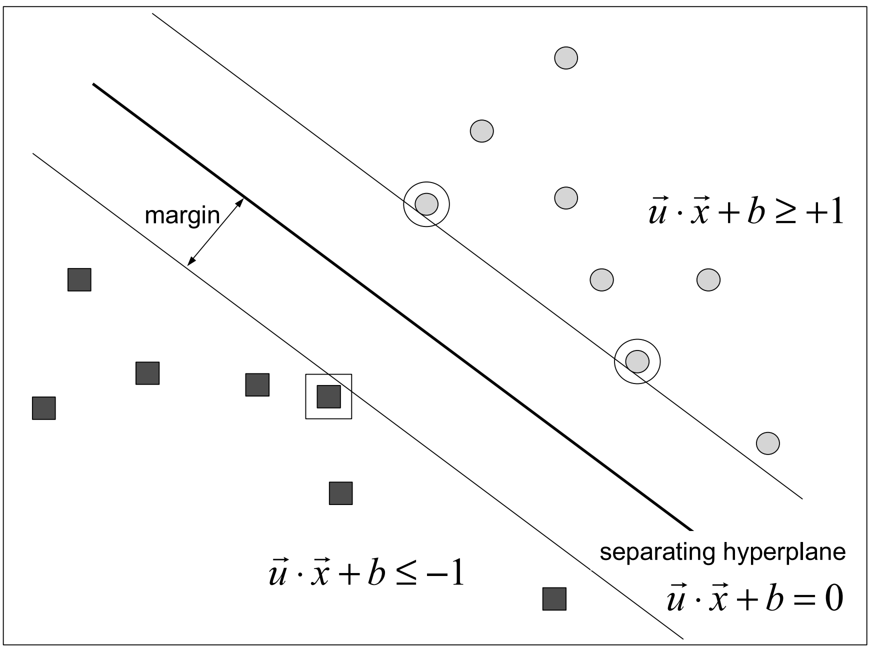


Figure 8.1. A decision hyperplane separates two classes (squares: y=-1 and circles: y=+1). The SVM algorithm seeks to maximise the margin around a hyperplane that separate a positive class from a negative class. The support vectors are framed.

A *single* support vector machine can only separate *two* classes: a positive class c_1 (indicated by $y = +1$) and a negative class c_2 (indicated by $y = -1$).

The SVM aims to find a class separating hyperplane with the largest possible margin, see Figure 8.1. (Vapnik, 1998; Joachims 2002) This results in a hyperplane which is defined by a normal vector $\vec{u}$ and an offset b.

In the classification step an unlabelled 'term frequency' vector is estimated to belong to the class

$$\hat{y} = \mathrm{sgn}(\vec{u} \cdot \vec{x}_i + b) \quad (2)$$

where $\vec{x}_i$ is the term frequency vector which represents document d_i. SVMs can be adjusted to different geometries in the feature space replacing the dot product by a kernel function $K(\vec{x}_i, \vec{x}_j)$. We have observed, however, that in the case of textual data the choice of the kernel function has a minimal effect on the accuracy of classification: Kernels which imply a high-dimensional feature space show slightly better results in terms of precision and recall, but they are subject to overfitting. (Leopold & Kindermann 2002)

The most important property of SVMs is that learning is independent of the dimensionality of the feature space. SVMs seek for the hyperplane that separates the two classes with the maximal margin (see Figure 8.1). Thus the separating hyperplane is defined in terms of data points touching the margin — the support vectors — rather than by coordinates of the feature space. This allows for a good generalisation even in the presence of a large number of features and makes SVM especially suitable for the classification of texts.

2.1.4 Decision Trees

Decision trees are classifiers which consist of a set of rules which are applied in a sequential way and finally yield a decision. They can be best explained by observing the training process, which starts with a comprehensive training set. It uses a divide and conquer strategy: For a training set M with labelled documents the word w_i or term is selected, which can predict the class of the documents in the best way. Then M is partitioned into two subsets, the subset M_i^+ with the documents in which w_i occurs, and the subset M_i^- with the documents without w_i. This procedure is recursively applied to M_i^+ and M_i^-. The procedure stops if all documents in a subset belong to the same class c_j generating a tree of rules with an assignment to actual classes in the leaves.

Decision trees are a standard tool in data mining (Mitchell, 1997). They are fast and scalable both in the number of variables and the size of the training set. For text mining, however, they have the drawback that the final decision depends only on relatively few terms. A decisive improvement may be achieved by boosting decision trees. This results in determining a set of complementary decision trees constructed in such a way that the overall error is reduced. Schapire and Singer (2000) use even simpler one step decision trees that contain only one rule and get impressive results.

2.2 Unsupervised Learning

Unsupervised learning methods aim at extracting all interesting patterns directly from the data. They do not require training data. In this section we describe clustering and reduction of dimensions.

2.2.1 Clustering

Clustering is one of the core data mining techniques. It refers to an unsupervised learning process in which individual items are grouped on the basis of their mutual similarity or distance. Again it is necessary to define implicitly or explicitly a similarity measure between documents. We may represent each document as a vector of word frequencies in some order and use the Euclidean distance, the cosine similarity (which has been defined in section 2.1.2) or some other similarity measure.

Clustering of documents has already been developed in the seventies (see, e.g., van Rijsbergen, 1979). A very simple clustering algorithm resembles the k-NN clustering in section 2.1.2. It includes documents d_i and d_j in the same cluster when the similarity measure $S(d_i, d_j)$ is below a given threshold.

Hierarchical clustering techniques create clusters by iteratively merging (agglomerative clustering) or splitting (divisive clustering) of previously identified clusters. The process of hierachical clustering therefore leads to the creation of a dendrogram, which is a tree of clusters, allowing one to adjust the clustering granularity according to the current needs.

Partitional techniques start with a fixed number of clusters, which are improved iteratively. A prominent member is the *k*-means algorithm (Hartigan, 1975). Its principle is that each cluster is represented by the means of all its members and serves as a basis for an iterative cluster regrouping. There exist a large number of clustering schemes, a survey is given by Hastie et al. (2001).

A specific variant is model based clustering, which assumes that the clusters are generated according to a statistical model. For text documents discrete distributions, e.g., multinomial distributions are most appropriate. This allows statistical techniques to estimate the most probable clustering and the adequacy of clusters (Nigam et al., 1999). The similarity measure is implicitly defined by the distributional model.

If there is no statistical model it is difficult to determine the optimal number and the validity of clusters. Potential cluster quality measures such as cluster stability, cluster compactness, or inter-cluster separation can be quantified with cluster validation indices.

2.2.2 Dimensionality Reduction

One key approach to data mining is the reduction of a large number of variables to a few constructs which capture the 'main' properties of the data. This is especially interesting for text mining where many thousand variables are common. We start with the term document matrix A in which each row contains the count of words of a document. *Principal component analysis* is a well-known approach from multivariate statistics (Hastie et al., 2001). It starts with the correlation matrix $A'A$ of all variables and uses eigen analysis to determine the largest eigenvectors of the correlation matrix. *Latent semantic analysis* (LSA) is a technique popular in text mining (Deerwester et al., 1990) which can be shown to yield the same results as principle component analysis (Thisted, 1988). A large number of comparable techniques has been discussed under the name of *factor analysis*.

The factors can be considered as independent linear combinations of the original variables that explain a maximal proportion of the variation in the dataset. In subsequent analyses, e.g., classifications or similarity computations, they may be used instead of the original variables without losing too much expressive power. The similarity of documents in terms of the principal components may be interpreted as topical similarity and can be used to find related documents, or documents matching some specified query. By this approach we may even estimate the similarity of documents even if they do not have any words in common. (Landauer and Dumais, 1997).

If different words have a high correlation to the same factor, this often indicates a similar meaning, i.e., synonymy. On the other hand the same word may have substantial correlations to two or more factors, which may indicate different meanings of the same word, i.e., polysemy.

One objection to latent semantic indexing is that it relies on the correlation matrix, and implicitly minimises square distances. More appropriate for count data in text mining is *probabilistic latent semantic analysis* (Hofman, 2001). It assumes a discrete unobservable variable z (latent factor) which may take the values 1 to m. The model assumes that for each word w_{ij} in a document d_i a value of the latent factor is generated according to a document specific distribution $p(z|d_i)$. Depending on the value of the latent factor the word w_{ij} then is generated according to a factor-specific distribution $p(w_{ij}|z)$.

The probabilities are estimated in an iterative way using the Expectation Maximisation (EM) algorithm which has been introduced by Dempster, Laird and Rubin (1977). Probabilistic latent semantic analysis (PLSA) results in a better linguistic interpretability and is compatible with the well corroborated linguistic models (see Chitashvili and Baayen 1993) of word frequency distributions.

LSA and PLSA have an interesting application to bibliometric search problems: The *Science Citation Index* offers a search by words or a combination of words. This means that documents that do not contain the word cannot be retrieved although they might deal with the requested content. Using LSA or PLSA queries and documents (or abstracts) could be mapped to its latent factors. This would enable a concept oriented search where synonyms of the word also indicate relevance to the query, and documents in which a word appears in a different meaning are rejected.

3. TECHNIQUES SPECIFIC TO TEXT MINING

As mentioned previously, text mining is data mining applied to natural language texts. The main issues which are connected with the transfer of general data mining techniques to the textual domain are the representation of texts, its pre-processing and the special statistical characteristics of textual data, which constitutes a special challenge to data mining algorithms.

Although the bag–of–words representation is very simple and effective, it neglects the succession of words in the texts and therefore abstracts away from the syntactic relations which exist between the different linguistic units. Furthermore, the level of analysis is not confined to words. Units that are smaller than words yield good results when small corpora are considered.

3.1 Morphological Pre-processing and Feature Selection

Preprocessing is concerned with the elimination of textual information which is irrelevant or even misleading to solving the subsequent data mining task. As a rule of thumb half of the words occur only once even in a large text corpus of some million running words (52% of the words of the corpus displayed in Figure 8.2 are hapax legomena). These words cannot occur in both the training set and the test set. They are therefore omitted.

Stemming is another pre-processing method in which the words that occur in the corpus are mapped to a basic form. Stemming is a more general notion than lemmatisation where the basic form is linguistically defined. The Porter Stemmer (Porter, 1980), which performs a cascade of regular expressions, is often used for English texts. Stemming, however, is a more challenging task for other languages that are more productive at the morpho-syntactic level.

Stemming can be performed at different levels of depth and has to be used with care. Resolving every morpho-syntactic rule assuredly leads to a loss of information. But even the removal of the plural morpheme can result in a loss of semantic information. Stricker et al. (2000) give an example for French: The word 'action' in the sentence "Le jugement est plus nuancé selon le domaine d'*action* du gouvernement." the word 'action' can be translated with action. However, in the sentence "Den Danske Bak a acquis en décembre dernier 90% des actions de Fokus Bank." the word 'actions' means shares, and this meaning is clearly indicated by the plural ending.

In some languages it is useful to split complex words such as, e.g., compounds into their morphological components, and preserve them for subsequent processing. The resulting features can be reduced further by applying other feature selection. Compound splitting can thus be considered as a sub-task of stemming, where compounds are split into its components. This is less necessary for the English language, but compound splitting is usually beneficial when applied to compounding languages like German or Dutch.

3.1.1 Term weighting and selection

Some words in the language's vocabulary are very frequent and equally distributed amongst the documents in the corpus. In many cases these words are superfluous from a statistical viewpoint and should be removed prior to the application of data mining algorithms. There are different techniques for performing this task.

The simplest method for the removal of uninformative words is to use a predefined list of stop words, and to delete all words in the text that match an element of the list. Stop word lists typically consist of function words (articles, pronouns, and conjunctions). The problem of stop word lists is that they may be inappropriate to the corpus or the task in question. In a corpus of texts on computers the word 'computer' will probably be equally distributed amongst the documents and thus fairly uninformative. In such a case the word 'computer' should be treated as a stop word. When the task is authorship recognition function words may be important cues for authorship recognition, although they are unlikely to be useful for content classification.

Other methods of identifying uninformative terms make direct use of its statistical distribution among the texts of the corpus. When pre-processing is performed prior to classification the distribution of terms in different classes in the training set can be compared against each other. Terms are omitted when a statistical test suggests that they are equally distributed in different classes. Pearson's chi-squared test, for instance, has been applied successfully (Paaß et al., 2002) prior to text classification based on sequences of syllables. Similar techniques such as information gain, mutual information, cross entropy or odds ratio are described in (Mladenic and Grobelnik, 1999).

Term weighting schemes like the well known *inverse document frequency* exploit the distribution of term frequencies in the text. The number of documents in which a given term appears is called *document frequency* denoted as *df*. The so called inverse document frequency, which was defined by (Salton, 1983) as $idf = (\log df)^{-1}$, is widely used in the literature on automatic text processing in order to tune term–frequencies according to the thematic relevance of a term. Other term weighting schemes such as, e.g., the redundancy used by (Leopold and Kindermann 2002) consider the entire distribution over the documents rather than solely the number of texts. A survey about different weighting schemes is given in (Manning and Schütze 1999). Although *idf* is the most popular weighting scheme, other indexing functions have also been used, including probabilistic techniques (Gövert et al., 1999). These weighting schemes are especially useful when the length of the documents exceeds some 1,000 words.

3.1.2 Statistic properties of textual units: Zipf's law

The skewness of frequency distribution makes linguistic data a special challenge for any statistical method of analysis. Zipf's law, which is empirically very well confirmed, ensures that the *r*-most frequent word in the corpus occurs $f(r) = a/(c+r)^g$ times, where *a*, *c* and *g* are parameters. (*g* varies from 1 corresponding to normal language use to 2 corresponding to a fairly restricted or standardised language use such as, e.g., in the Reuters news wire corpus.). Interestingly Zipf-like distributions hold for nearly any type of linguistic units: frequencies of syllables or lemmas also follow the Zipf distribution. As a rule of thumb the parameter *g* decreases with increasing unit size. (Note that double logarithmic scaling in Figure 8.2 displays functions of the form $f(r) = a/r^g,\ a, r > 0$ as a straight line with negative slope.) One consequence of Zipf's law is that the word frequency distribution is very skewed, i.e., some few words are very frequent (some 10^5 occurrences), whereas the frequency of most of the words is some magnitudes smaller (< 10). Usually (i.e., unrestricted language use assumed) half of the words in a text corpus occur only once. Unfortunately it is well known that especially the rare words are particularly informative about the content of the document in which they occur. This means that rare words may not be considered irrelevant and may not be omitted prior to the text mining process.

A further consequence of Zipf's law is that most of the words in the corpus are absent in most of the documents in the corpus. This phenomenon is usually addressed as the sparse data problem. The so called bag–of–words representation of documents, which counts the number of occurrences of each word in the document and thus ignores the succession of words, is often used for text mining purposes. Text corpora often contain some millions of running words and some 100,000 different word types, and most of the types which occur in the corpus do not occur in a particular document. Therefore vectors that result from the bag–of–word representation are sparse (i.e., most entries equal zero) and moreover they are very long (some 100,000 dimensions).

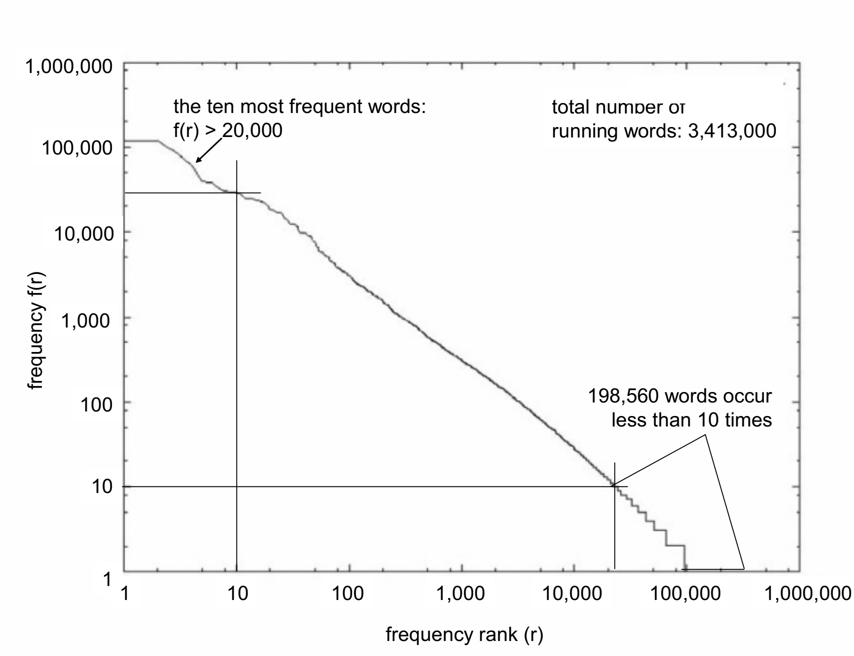


Figure 8.2. Rank-Frequency Distribution of Frankfurter Rundschau, July 1998

3.1.3 Subword Units

In order to relieve the sparse data problem it is sometimes useful to use units that are smaller than words, so called sub-word units (sequences of letters, or syllables) instead of words. Sub-word units, yield good results especially when small corpora are considered. The F-scores presented in Figure 8.3 were obtained by a SVM classifier from a corpus of German radio programs. The corpus consists of 950 documents comprising about 650 running words each. The texts were converted to a phonetic transcript using the BOSS II speech synthesis system (Stöber et al., 2000). In the corpus the words consist on average of 5.3 phonemes, and syllables comprise on average 2.8 phonemes. We experimented with different kinds of units: sequences of 2 to 6 phonemes (phoneme-n-grams), sequences of 1 to 6 syllables (syllable-n-grams) and 1 to 3 words.

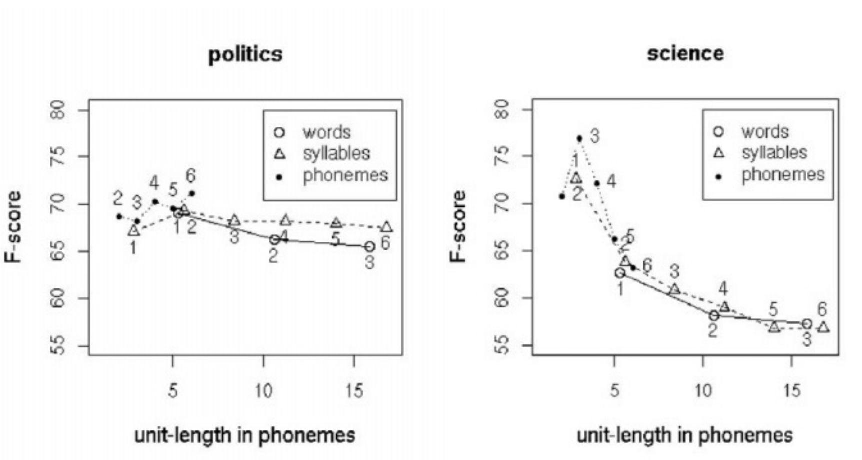


Figure 8.3. Classification accuracy achieved with different representations of texts. Left panel: large class (220 documents); right panel: medium size class (120 documents).

In Figure 8.3 the F-scores are presented for the largest class in the corpus, 'politics', which comprises 220 documents, and for a smaller class, 'science' comprising 120 documents. One can see that accuracy of SVM Classifier depends strongly on the size of the units. The optimal unit size is smaller when the class contains fewer documents (3 phonemes for the class 'science'). The large class makes little profit from small unit sizes.

An explanation for these results is the well known fact that there is a trade off between the length of linguistic units and their frequency. This has already been proven for words (Guiter, 1974), but it also applies to other units like sequences of syllables. This means that the smaller the categories and the longer the units (types) the more types become so improbable that they exclusively occur in either the test set or the trainings set. Therefore units have to be shorter in order to compensate for the small size of the class.

3.1.4 Shallow parsing

The notion Natural Language Parsing addresses the task of automatically detecting the linguistic structure of the sentences of a text. Although the above mentioned bag–of–words representation yields surprisingly good results for many text mining applications, it seems obvious that subsequent algorithms work the better the more structure is extracted from the original texts. However, even if it were possible to formalise and represent the complete grammatical and lexical structure of a natural language, the parser, would still need a very high degree of robustness and efficiency. Realising such a system for large amounts of texts is impossible for the time being. This has led to so called shallow parsing approaches, in which certain language regularities which are known to cause complexity problems are handled in a pragmatic way (Neumann & Piskorski, 2002).

Neumann and Schmeier (2002) showed, for example, that morphogical analysis of short German texts seems to be a better choice than simple trigramming. For the collection of small documents (emails of 60 words in average) SVM yielded the best results when combined with a shallow parsing compared to trigrams or morphs (accuracy 61.42 vs. 58.29 and 54.29 respectively). For the collection of longer texts (average length 578 words) there was no significant difference in accuracy between morphs and trigrams.

3.2 Presentation of Results

3.2.1 Graphical representation

Self-organising Maps (SOM) were invented in the early 80s (Kohonen, 1980). They use a specific neural network architecture to perform a recursive regression leading to a reduction of the dimension of the data. For practical applications SOMs can be considered as a distance preserving mapping from a more than three-dimensional space to two-dimensions. A description of the SOM algorithm and a thorough discussion of the topic is given by Kohonen (1995).

Figure 8.4 shows an example of a SOM visualising the semantic relations of news messages. First the news messages were represented in a four-dimensional semantic space, where each coordinate corresponds to a topic category ('culture', 'economy', 'politics' and 'sports'). Then the SOM algorithm is applied (with 100×100 nodes using Euclidean metric) in order to map the four-dimensional document representations to two dimensions admitting a minimum distortion of the distances. The grey tone indicates the topic category.

3.2.2 Text summarisation

Visualisations described above have to be accompanied by some description providing the user with some additional information about the content of the respective documents. One of the options for providing this additional information is text summarisation. Text transformation can be defined as a reductive transformation of a source text by the selection and generalisation of what is important in the source (Spark-Jones, 1999).

In the currently most explored approach for summarisation the summary is composed of sentences which are selected as semantically representative for the document content. An example of such an extractive multi-document summarisation approach is given by Kraaij et al. (2002).

Other recently addressed text summarisation research topics have been multi-lingual summarisation and hybrid multi-source summarisation (Chen, 2002). The potential of concepts and conceptual relations as a vehicle for terminological knowledge representation has been exploited by the knowledge–based text summarisation approach (Hahn & Reimer, 1999).

4. CASE STUDIES

4.1 Case study 1: Classification of Web Documents

Web mining is the use of data mining techniques to automatically discover and extract information from Web documents and services (Kosala & Blockeel 2000). Three main tasks may be solved by web mining: *Web structure mining* tries to discover regularities underlying the link structures in the web. *Web usage mining* evaluates the frequency and temporal sequence in the actual access of web pages by users. *Web content mining* describes the discovery of useful information from web documents.

The EU project Diastasis aims at describing the Web surfing behaviour of citizens. The ultimate goal is to establish representative statistics which, on the one hand, describe the socio-economic situation of people, and, on the other hand, the content of the web pages they download. This information will be compiled and distributed by national statistical institutes.

A representative sample of people, who have given their consent, is observed with respect to their web surfing behaviour. This means that the IP addresses of web pages they download as well as the exact time of download are recorded in a log file.

In order to describe the content of web pages a hierarchy of meaningful categories is required, which is expressive enough to cover all aspects of web pages. We used the Yahoo web directory (http://www.yahoo.com) for this purpose, as each category is linked to example pages which can be used for training purposes. The idea is to train a text classifier for the higher categories of the hierarchy. This offers the opportunity to classify the content of arbitrary web pages into meaningful categories.

Initially each linked HTML document from the directory is downloaded. Images and HTML code are deleted so that only text remains. This text is pre-processed with the usual pre-processing steps (stemming, stopword removal, etc.) described above. Finally SVM classifiers are trained for each category of interest.

To describe the content of web pages used by the citizens the corresponding pages have to be downloaded again according to their IP addresses. Subsequently these pages may be classified into one or more categories of the ontology.

A difficulty arises because the documents contain different languages. For the specific population of Barcelona the main languages are Spanish, Catalan and English. To solve this problem a language classifier is used which assigns documents to languages based on triples of letters. Yahoo directories exist for all three languages and the directories have nearly identical structure. Hence classifiers can be trained for the Yahoo hierarchy in these three languages.

To identify the users's surfing behaviour the temporal sequence of page views has to monitored. To accommodate this user sessions are identified. A session is defined as a continuous browsing period containing no idle periods longer than thirty minutes of length. It is non-trivial to create a completely accurate session for a user, considering the web log mechanism. For example, when a user accesses a locally cached web resource this access does not appear in the log; thus repeated visits to sites will often only appear once. Our method of recreating sessions via external observation is also susceptible to error. If a user spends thirty minutes reading a web page before visiting the next page in the logical session, they will be identified as two separate sessions.

Dynamic web sites also present a challenge. Here a page is dynamically constructed by the server and may change after each visit; for example, most popular portal web sites. Therefore it is difficult to build an accurate classification model of the site's pages, as an access after some time may yield completely different content. These difficulties can be overcome with a more extensive logging system and real time training and classification.

To combine the web content data with the user data we first have to describe the temporal sequence of web utilisation, e.g., accessing financial information in the morning and visiting web auctions in the evening. This is a task of web usage mining. The resulting web surfing information is combined with socio-economic user data collected by questionnaires and surveys. The final information will be published by national statistical institutes.

The method outlined here can also be used for bibliographical purposes. Publications, possibly in different languages, are first classified according to their language and then inserted into different topic categories, which might be defined, for instance, by abstracts and their classification codes of the *Science Citation Index* (instead of the documents of the Yahoo web directory).

4.2 Case study 2: Authorship Attribution

Authorship attribution can be considered as a special case of text classification, in the sense that a text is classified according to whether it was written by a specified author or not. However various approaches to authorship attribution differ significantly from usual text classification techniques.

There are a number of statistical techniques which have been imported from other fields and which dominate the field of computer–based authorship attribution. Most notably are the Efron-Thisted Test (Thisted & Efron, 1987), QSUM (or cusum) (Holmes, 1998) feed–forward artificial neural networks (Tweedie et al., 1996), Radial Basis Function (RBF) networks (Lowe and Matthews, 1995), genetic algorithms (Holmes and Forsyth, 1995), Recurrent neural networks (Towsey et al., 1998). According to Rudman (1998) approximately 1,000 style markers have already been isolated for authorship attribution. There is, however, no agreement of significant style markers amongst researchers

Authorship attribution has previously suffered from the problem that the important features in a document are unknown and that a text as a whole cannot be analysed. The use of a limited set of function words or ‘short words’ is clearly restrictive and there is an ongoing discussion on the relevance of appropriate style marker.

This calls for the application of techniques such as support vector machines which allow one to process bag–of–words representations of complete texts rather than a small number of selected features. SVMs can process documents of significant length, databases with a large number of texts and do not require pre-determined features. As Leopold and Kindermann (2002) show, SVM are capable of managing input vectors with a very large number of dimensions (up to 400,000), with no term selection required.

For the experiments on authorship attribution data from the Berliner Zeitung (BZ), a daily newspaper in Berlin, were used. We used the data from December 1998 to February 1999. The articles are divided into twelve topics. From the three largest topics: politics (1,200 articles); economy (550 articles); and local affairs (3,233 articles) all articles with more than 200 words were considered. The resulting corpus consisted of 2,652 documents, about 1,900,000 running words (tokens) and about 120,000 different words (types). These documents were represented in two different ways:

1. A vector word counts was generated from the document. Neither stemming nor stop word removal was performed. (a simple bag–of–words representation);
2. For each document tagwords were extracted and bigrams were generated from them. Additionally the number of words with a given word length is counted. The document is represented by the vector of counts of tagwords, of bigrams, and of word lengths.

An SVM classifier was trained for seven authors in the corpus. Different parameter settings have been applied and five-fold cross-validation was performed. Tables 8.1 and 8.2 show the results for the optimal choice of parameters (Diederich et al., 2003).

Table 8.1. Results of classification based on word forms

name of author	*target author*		*other authors*		*percent*	
	# correct	*# false*	*# correct*	*# false*	*precision*	*recall*
Aulich	94	14	2652	0	100.0	87.0
Fuchs	98	20	2642	0	100.0	83.1
Kunert	71	29	2659	1	98.6	71.0
Muenner	80	7	2673	0	100.0	92.0
Neumann	73	38	2647	2	97.3	65.8
Schmidl	66	28	2666	0	100.0	70.2
Schomaker	25	57	2678	0	100.0	30.5

In most cases the target author was recognised correctly. Notably the number of false negative classifications is extremely low. That is erroneously attributed authorship is very improbable. There is no significant difference between the results for words (Table 8.1) and tagwords (Table *8.2*). This suggests that SVM combined with the bag–of–words representation in its simplest form is sufficient for a reliable identification of authorship.

Table 8.2. Results of classification based on tagwords, bigrams of tagwords, and word lengths

name of author	*target author*		*other authors*		*percent*	
	# correct	*# false*	*# correct*	*# false*	*precision*	*recall*
Aulich	85	23	2652	0	100.0	79.0
Fuchs	89	29	2642	0	100.0	75.0
Kunert	61	39	2660	0	100.0	61.0
Muenner	67	20	2673	0	100.0	77.0
Neumann	51	60	2649	0	100.0	46.0
Schmidl	60	34	2666	0	100.0	63.8
Schomaker	17	65	2677	1	94.4	21.0

Authorship attribution using SVMs allows for the verification of a pretended authorship given that enough training examples that is, previous publications of the author in question are available. Note that precision is very high in the results presented above. This means that if the classifier is able to identify an author its decision is very reliable.

4.3 Case study 3: Classifier Induced Semantic Spaces

Latent Semantic Analysis (Landauer and Dumais, 1997) as well as Probabilistic Latent Semantic Analysis (Hofman, 2001), which is described in section 2.2.2, are often used for the construction of semantic spaces. Semantic spaces typically reflect some aspect semantic nearness of linguistic units. The dimensions of such a semantic space are often interpreted as 'artificial concepts' which represent common meaning components of different words and documents. Such artificial concepts, however, cannot be interpreted in a semantically transparent way.

Another way of generating semantic spaces which produce a semantically transparent representation can be constructed from the internal representation of supervised text classifiers. Recall the classification step of Support Vector Machines, described in section 2.1.3. Estimating the class membership by equation (2) consists of a loss of information since only the algebraic sign of the right hand term is evaluated. However, the value of

$$v = \vec{u} \cdot \vec{x} + b$$

in equation (2) is a real number and can be used in order to create a real valued semantic space, rather than just to estimate if $\vec{x}$ belongs to a given class or not.

Suppose there are several, say K, classes of documents. Each document is represented by an input vector $\vec{x}$. For each document the variable $y_i^k \in \{-1,+1\}$ indicates whether $\vec{x}$ belongs to the *k*th class, $k = 1,\dots,K$, or not. For each class an SVM can be trained which yields the parameters $\vec{u}^k$ and b^k. After the SVMs have been learned, the classification step (equation (2)) can be applied to a (possibly unlabeled) document represented by $\vec{x}$ resulting in a *K*-dimensional vector $\vec{v}$, where the *k*th component is given by

$$v^k = \vec{u}^k \cdot \vec{x} + b^k$$

The component v^k quantifies how much a document belongs to class *k*. Thus the document represented by its term frequency vector is mapped to the *K*-dimensional vector in the classifier induced semantic space. Each dimension in this space can be interpreted as the membership degree of the document to each of the *K* classes.

Figure 8.4 shows a Self-organising Map (see section 3.2.1), which is generated from a classifier induced semantic space. SVMs for the four classes ‘culture’, ‘economy’, ‘politics’, and ‘sports’ were trained by news messages from the ‘Basisdienst’ of the German Press Agency (dpa) April 2000. Classification and generation of the SOM was performed for the news messages of the first 10 days of April. 50 messages were selected at random and displayed as white crosses. The categories are indicated by different grey tone. Shadings within the categories indicate the confidence of the estimated class membership.

It can be seen that the change from sports (15) to economy (04) is filled by documents which cannot be assigned confidently to either classes. The area between politics (11) and economy (04), however, contains documents, which definitely belong to both classes. Note that classifier induced semantic spaces go beyond a mere extrapolation of the annotations found in the training corpus. It gives an insight into how typical a certain document is for each of the classes. Furthermore Classifier induced semantic spaces allow one to reveal previously unseen relationships between classes. The bright islands in area 11 on Figure 8.4 show, for example, that there are messages classified as economy which surely belong to politics.

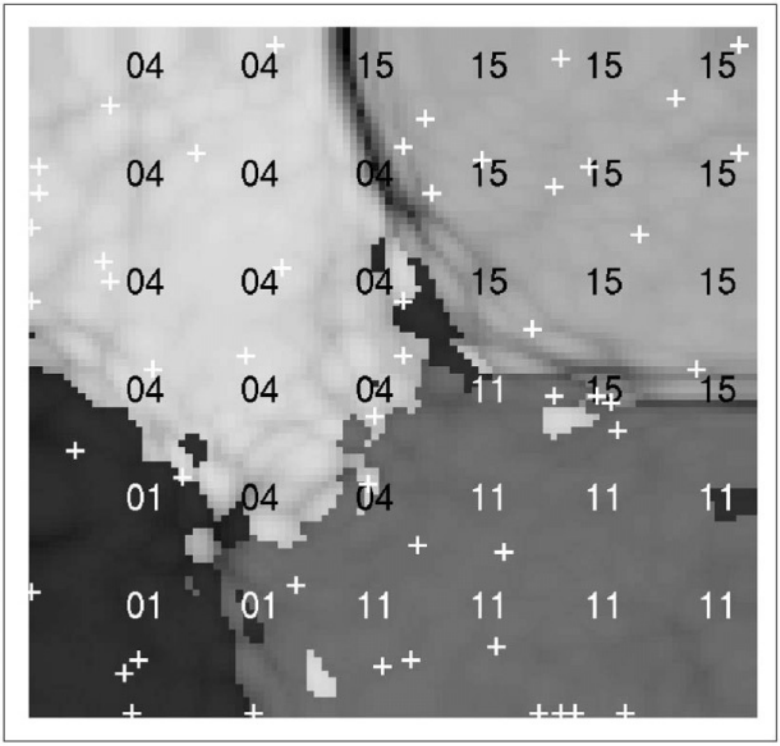


Figure 8.4. SOM of a classifier induced semantic space. The numbers indicate the classified topics 01: culture, 04: economy, 11: politics, 15: sports.

Classifier induced semantic spaces can also be used for bibliometric purposes such as, e.g., a refinement of the coarse–grained classification codes of the *Science Citation Index* or the identification of emerging topics; just substitute the topic classes by the respective classification codes and train SVMs by the SCI abstracts.

– *Refinement of search*: When a document collection is represented in a SOM like the one in Figure 8.4 one can see how much a document belongs to different classes. A document which is located at the border between two classes (say 'sports' and 'art') is likely to belong to both classes (perhaps they deal of dance) although they might be classified into only one class. Representing abstracts in a SOM allows for searching exactly these borderline documents.

- *Detection of subcategories:* Clusters which can be observed on a SOM are likely to belong to a 'sub-category' which is not explicitly represented in the coding scheme of the *Science Citation Index*. This results a more fine–grained categorisation of the documents collection. The thematic domain of the new subcategory can be inferred from the position on the SOM, since the dimensions of the semantic space are semantically interpretable. *Clustering* of unlabelled documents according to their similarity can be deployed for the detection of related information.
- *Emerging topics:* Publications which represent unusual themes are likely to be separated from all other documents in the space. Items which do not fit to the category codes are situated in the negative simplex of the coordinate system. When publications in the negative simplex $v^k < 0,\ k = 1,\ldots,K$, accumulate it is likely that a 'new' topic has emerged, which is not covered by the present classification codes.

5. CONCLUSION

Text mining offers a variety of methods for the automatic analysis of texts, which can also be gainfully applied to bibliometric problems and patent statistics. Latent Semantic Analysis and the more recently developed Probabilistic Latent Semantic Analysis allows for a concept oriented rather than key-word based search in bibliographical indices or patent databases. The classification of publications or abstracts into a predefined ontology like the one defined by the classification codes of the SCI can be done automatically using document classification techniques. Authorship attribution can be used for the detection of plagiarism. Classifier induced spaces can be utilised for the identification of emerging topics. Self–organising Maps can help to detect subcategories. They can also be used for a refined search in a collection of categorised abstracts or documents.

REFERENCES

Andrews, R., Geva S. (1994). *Rule extraction from a constrained error backpropagation MLP*. Australian Conference on Neural Networks, Brisbane, Queensland 1994 (pp. 9–12).

Baeza-Yates, R., Ribeiro-Neto, B. (1999). *Modern information retrieval*. Addison-Wesley.

Chen, H.H. (2002). *Multilingual summarization and question answering*. Workshop on Multilingual Summarization and Question Answering, COLING'02, Taipeh, Taiwan 2002.

Chitashvili, R.J., Baayen, R.H. (1993). *Word frequency distributions*. In G. Altmann, L. Hřebíček (Eds.), Quantitative Text Analysis (pp. 54–135). Wvt: Trier.

Deerwester, S., Dumais, S.T., Landauer, T.K., Furnas, G.W., Harshman, R.A. (1990). Indexing by latent semantic analysis. *Journal of the American Society of Information Science,* 41 (6), 391–407.

Dempster, A.P., Laird, N.M., Rubin, D.B. (1997). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society,* B, 39, 1–38.

Diederich, J., Kindermann, J., Leopold, E., Paaß, G. (2003). Authorship attribution with Support Vector Machines. *Applied Intelligence,* 19 (1–2), 109–123.

Dumais, S., Platt, J., Heckerman, D., Sahami, M. (1998). *Inductive learning algorithms and representations for text categorization*. In Proceedings of the 7th International Conference on Information and Knowledge Management (pp. 148–155). ACM.

Gövert, B., Lalmas, M., Fuhr, N. (1999). *A probabilistic description–oriented approach for categorising Web documents*. In Proceedings of CIKM-99, 8th ACM International Conference on Information and Knowledge Management, Kansas City, Missouri, 1999 (pp. 475–482). ACM.

Guiter, H. (1974). *Les rélations fréquence – longueur – sens des mots (langues romanes et anglais),* In XIV congresso internazionale di linguistica e filologia romanza (pp. 373–381). Napoli.

Hahn, U., Reimer, U. (1999). *Knowledge–based text summarization*. In: I. Mani, M. T. Maybury (Eds.), Advances in Automated Text Summarization (pp. 215–232). Cambridge, London: MIT-Press.

Hand, D., Mannila, H., Smyth, P (2001). *Principles of data mining*. MIT Press.

Hartigan, J.A. (1975). *Clustering algorithms.* New York: John Wiley.

Hastie T., Tibshirani, R., Friedman, J. (2001). *The elements of statistical learning.* New York: Springer.

Hofman, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning,* 42, 177–196.

Holmes, D.I. (1998). The evolution of stylometry in Humanities Scholarship. *Literary and Linguistic Computing,* 13 (3), 111–117.

Holmes, D.I., Forsyth, R.S. (1995). The Federalist revisited: New directions in authorship attribution. *Literary and Linguistic Computing,* 10 (2), 111–127.

Kohonen, T. (1980). *Content–adressable memories.* Springer.

Kohonen, T. (1995). *Self-organising Maps.* Springer.

Kosala, R. Blockeel, H. (2000). *Web mining research: A Survey*. In P.S. Bradley, S. Sarawagi, U.M. Fayyad (Eds.), SIGKDD Explorations: Newsletter of the Special Interest Group (SIG) on Knowledge Discovery & Data Mining, ACM, 2 (pp. 1–15). ACM Press.

Kraaij, W., Spitters, M., Hulth, A. (2002). *Headline extraction based on a combination of uni- and multidocument summarization techniques.* In Proceedings of the ACL workshop on Automatic Summarization/Document Understanding Conference DUC 2002 , June 2002, Philadelphia, USA.

Joachims, T. (1998a). *Making large-scale SVM learning practical*, Technical report University of Dortmund.

Joachims, T. (1998b). *Text categorization with Support Vector Machines: learning with many relevant features*. Proceedings of the 10th European Conference on Machine Learning, Springer Lecture Notes in Computer Science, Vol. 1398 (pp. 137–142). Springer.

Landauer, T.K., Dumais, S.T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104 (2), 211–240.

Lang, K. (1995). *Newsweeder: Learning to filter netnews.* In A. Prieditis, S. Russell (Eds.), Proceedings of the 12th International Conferrence on Machine Learning (pp. 331–339). San Francisco: Morgan Kaufmann Publishers.

Leopold, E., Kindermann, J. (2002). Text categorization with Support Vector Machines. How to represent texts in input space? *Machine Learning,* 46, 423–444.

Lowe, D., Matthews, R. (1995). Shakespeare vs. Fletcher: A stylometric analysis by radial basis functions. *Computers and the Humanities,* 29, 449–461.

Manning, C.D., Schütze, H.(1999). *Foundations of statistical natural language processing*. Cambridge MA, London: MIT Press.

Mitchell, Tom (1997). *Machine Learning*. Boston et al.: McGraw-Hill.

Mladenic, D., Grobelnik M. (1999). *Feature selection for unbalanced class distribution and naive Bayes*. In I. Bratko, S. Dzeroski (Eds.), Proceedings of the Sixteenth International Conference on Machine Learning (ICML 1999) (pp. 258–267). San Francisco: Morgan Kaufmann.

Neumann, G., Schmeier, S. (2002). Shallow natural language technology and text mining. *Künstliche Intelligenz*, 2002 (2), 23–26.

Neumann, G., Piskorski, J. (2002). A Shallow text processing core engine. *Computational Intelligence*, 18 (3), 451–476.

Nigam, K., McCallum, A.K., Thrun, S., Mitchel, T. (1999). Text classification from labeled and unlabeled documents using EM. *Machine Learning*, 39 (1/2), 103–134.

Paaß, G., Leopold, E., Larson, M., Kindermann, J., Eickeler, S. (2002). *SVM Classification using sequences of phonemes and syllables.* Tapio Elomaa & Heikki Mannila & Hannu Toivonen (Eds.), Proceedings of the 6th European Conference on Principles of Data Mining and Knowledge Discovery (PKDD 2002); August 19–23, 2002 Helsinki, Finland, Lecture Notes in Artificial Intelligence 2431 (pp. 373–384) Berlin, Heidelberg: Springer.

Porter, M.F. (1980). An algorithm for suffix stripping. *Program (Automated Library and Information Systems)*, 14 (3), 130–137.

Rudman, J. (1998). The state of authorship attribution studies: some problems and solutions. *Computers and the Humanities,* 31, 351–365.

Salton, G., McGill, M.J. 1983. *Introduction to modern information retrieval.* New York: McGraw Hill.

Shapire, R.E., Singer, Y. (2000). BoosTexter: a boosting based system for text categorization. *Machine Learning*, 39, 135–168.

Sparck-Jones, K. (1999). *Automatic summarizing: factors and directions.* In I. Mani, M.T. Maybury (Eds.), Advances in Automated Text Summarization.

Srivastava, J., Cooley, R., Deshpande, M., Tan, P.-N. (2000). Web usage mining: discovery and applications of usage patterns from web data, *SIGKDD Exploratins,* 1 (2), 12–23.

Stöber, K., Wagner, P., Helbit, J., Köster, S., Stall, D., Thomae, M., Blauert, J., Hess, W., Hoffmann, R., Mangold, H. (2000). *Speech synthesis by multilevel selection and concatenation of units from large speech Corpora.* In: W. Wahlster (Ed.), Verb-mobil. Springer, 2000.

Stricker, M., Vichot, F., Dreyfus, G., Wolinski, F. (2000). *Vers la conception de filtres d'informations efficaces*. In Reconnaissance des Formes et Intelligence Artificielle (RFIA '2000) (pp. 129–137).

Thisted, R., Efron, B. (1987). Did Shakespeare write a newly discovered poem? *Biometrika*, 74 (3), 445–55.

Thisted, R. (1988). *Elements of statistical computing.* London: Chapman&Hall.

Towsey, M., Diederich, J., Schellhammer, I., Chalup, S., Brugman, C. (1998). Natural language learning by recurrent neural networks: A comparison with probabilistic approaches. *Computational natural language learning conference*. Australian Natural Language Processing Fortnight. Sydney: Macquarie University, 15–17 Jan 1998.

Tweedie, F.J., Singh, S., Holmes, D.I. (1996). Neural network applications in stylometry: the federalist paper. *Computers and the Humanities,* 30, 1–10.

van Rijsbergen, C.J. (1979). *Information Retrieval.* London, Boston: Butterworths.

Vapnik, V.N. (1998). *Statistical Learning Theory.* New York et al.: Wiley & Sons.

Weiss, S.M., Apt, C., Damerau, F., Johnson, D.E., Oles, F.J., Goetz, T., Hampp, T. (1999). Maximizing textmining performance. *IEEE Intelligent Systems*, 14 (4), 63–69.

Chapter 9

OPENING THE BLACK BOX

Analytical Approaches and their Impact on the Outcome of Statistical Patent Analyses

Sybille Hinze and Ulrich Schmoch

Fraunhofer Institute for Systems and Innovation Research, Karlsruhe, Germany
E-mail: s.hinze@isi.fraunhofer.de

Abstract: The paper presents methodological approaches to statistical patent analyses. The references to timescales, countries of origin, and patent offices largely determine the outcome of such analyses; in particular, for country comparisons. For instance, considerable differences appear if results are compared based on priority, application, or grant years. For interpreting the patent figures at specific offices it proves important to consider the geo-strategic position of the office. Advanced approaches such as the triad concept lead to more balanced results, but their assessment has to include a consideration of international patent flows. For quality indicators it has to be taken into account that patents are always two-dimensional and have technical and economic aspects. In principle early quality indicators primarily cover the technological content. A further issue is the definition of samples aiming at a large number of suitable documents which is sometimes contradictory to the target of completeness.

1. INTRODUCTION

Since the seminal work of Schmookler (1966) on the link between technical inventions and economic growth, some first attempts have been made to use patent indicators to assess technological performance (see for example Kuntze et al., 1975, or Campbell and Nieves, 1979). The real breakthrough occurred in the middle of the eighties with the provision of and public access to electronic patent databases, allowing statistical analyses of large data sets (see, for instance, Narin et al., 1987, or Gerstenberger, 1992). In contrast to bibliometrics, the methodological discussion of patent indicators is limited to a few publications (for example Basberg, 1987,

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 215-235.

Pavitt, 1985, Schmoch et al., 1988, Archibugi, 1992, Griliches, 1990, or OECD, 1994, Schmoch 1999a). However, a flood of papers using patent indicators has been published, often without a precisely defined methodology. An annoying consequence of this practice is the publication of patent analyses of topical issues with differing, sometimes contradictory results. As the readers of such studies do not have a detailed insight into the methodology used, the results of patent statistics seem to be largely arbitrary, so that the methods based on patent statistics prove to be a black box with electronic patent data as the input and statistical tables as the output. In this paper, we try to open this black box and to demonstrate the consequences of specific methodological approaches. We will suggest preferred approaches in order to contribute to a broader agreement on the basic standards of patent statistics and thus to improve their credibility.

2. APPROPRIATE REFERENCES

In a first step we will discuss the use of different references and the effect they have on the results of patent analyses. The factors concerned and discussed in the following sections are: the timescale; the country of origin of a patent; and the patent office chosen as a basis of analysis. The differences will be shown which arise from the selections made and the specific methodological procedures. In particular, we want to raise the awareness and increase the sensibility of the use of data, the difficulties regarding their interpretation, and the dangers of misinterpretation.

2.1 References to Timescales

Patents are not applied for and then granted at the same time. In general there is a significant delay between application and grant. For instance, at the *European Patent Office* (EPO) the granting procedure took more than 49 months for about half the patents granted in 2002, after 70 months[1] the granting procedure was only completed for about 80% of the patents (EPO, 2002). The time lag between patent application and grant underlines the importance of choosing an appropriate timescale as the basis for patent analyses. At the same time it is a strong argument in favour of using the application date or, even more precisely, the priority date as the time reference in order to be as topical as possible.

[1] The application date at the EPO was taken as the time reference, based on priority data, the time delay would be even greater at about 82 months.

The priority date is the date of the first application for a certain patent and thus the date that is most closely related to when the invention covered by this patent was made. If patent indicators are used to reflect the applied R&D activities of firms or research institutions, the priority date proves to be the most appropriate reference.

The priority date is an important point in time in the legal grant procedure, as it determines the reference for the requirement of patentability. However, it is possible to modify the basic patent within the first year after application, the so-called 'priority year'. Then the basic patent is combined with the additional new features. In this case an additional priority date is assigned to this specific document which refers to the introduction of the new features, while the original priority date remains valid for the original content. As a consequence more than one priority date may appear in a patent application with a significant impact on the patent statistics. In order to avoid double counting of patents owed to multiple priority dates it is recommended to count only the first (or oldest) priority date given, as this is the date closest to the date of the invention.

Table 9.1. Number of patent applications at the EPO in the priority year 2000 for selected countries by different search methods

Inventor country	*Priority year*		
	2000	*2000 excluding 1999*	*2000 excluding 1999 and 1998*
USA	39,802	29,765	29,765
Germany	26,969	22,536	22,536
Japan	25,530	20,533	20,533

Source: PATDPA, Host STN

The effects of the multiple assignment of priority dates are made visible in Table 9.1. The significantly higher number of patent applications when searching for the priority year 2000 indicates double counting owed to multiple priority years in the patent application. It is quite easy to avoid this effect by simply excluding the year previous to the year considered. This results in the patent applications containing more than one time reference being counted only once. As amendments can only be introduced in the first year after the basic priority date, it is sufficient to exclude one year; the exclusion of two years does not have an additional effect (see Table 9.1). However, where the US Patent and Trademark Office (USPTO) is concerned we recommend excluding at least two years owing to the different framework of legislation in the US in comparison to Europe[2].

The date of first publication is a possible alternative time reference for patent analysis. In Europe, but also in Japan and Canada, patent applications are published strictly 18 months after the first priority date. This may be interpreted as the time when competitors become aware of an invention, so it has a direct impact on the development of technology. Nevertheless, we prefer the priority date as the reference owing to its close link with the finalisation of the invention and thus to R&D. If the grant date is used as a reference, the link to other indicators is even more difficult, as a specific grant year is related to applications with widely varying priority years. In addition, changes in grant figures are sometimes owing to administrative changes at the patent offices, for instance, rationalisation effects of information technology, or an increase of staff.

For many years it was not possible to analyse patent applications at the USPTO statistically, because only patents granted were made publicly available. Following the revision of the American Inventors Protection Act in 1999, the USPTO has also started to publish patent applications after eighteen months. Owing to the former practise of only publishing granted patents, USPTO–based patent analyses lacked up-to-dateness owing to the time delay between patent application and grant. Since 2000 the USPTO has also been publishing patent applications 18 months after the priority date, but there are certain exceptions which allow firms to avoid prompt publication of their inventions. Upon request, patent applications are not published at this time if they are not intended to be applied for in other countries. If at a later stage the firm reconsiders its application strategy and decides to file a foreign patent, it has to notify the USPTO and the respective patent application will be published. If this notification is not given, the respective patent application can be declared null and void (Department of Commerce, 2000). Owing to this new legislation we can expect at least a large share of US patent applications to be available for early statistical analysis.

The consequences of this new situation are illustrated in Figure 9.1 which depicts the trends in applications and grants differentiated by US and non-US inventors. Concerning grants, we see a substantial decrease for US inventors since the priority year 2000 owing to the time lag between application and grant. For non-US inventors the drop already begins in 1998 because of the delayed entry of foreign applications into the US system linked to the priority year for foreign applications and the further delay of

2 Japan and Canada have patent legislation similar to European legislations.

PCT[3] applications with US designation. The US applications reach a significant number in 2001, whereas the year 2002 is not complete yet owing to the publication delay of 18 months, similar to the European legislation. Two reasons explain the peak in applications by non-US inventors in 2000. First, the USPTO application year 2000 was the first year the new legislation was valid, which, from the perspective of non-US inventors, is generally the priority year 1999 so that these applications could be published earlier. Second, many non-US applications follow the PCT route and enter the US system with a substantial delay[4]. This may explain why the priority year 2001 is not complete yet. In any case, the early publication of applications at the USPTO offers new possibilities for statistical analyses. However, we need further information about the shares of early and late published applications to enable appropriate interpretation of the figures.

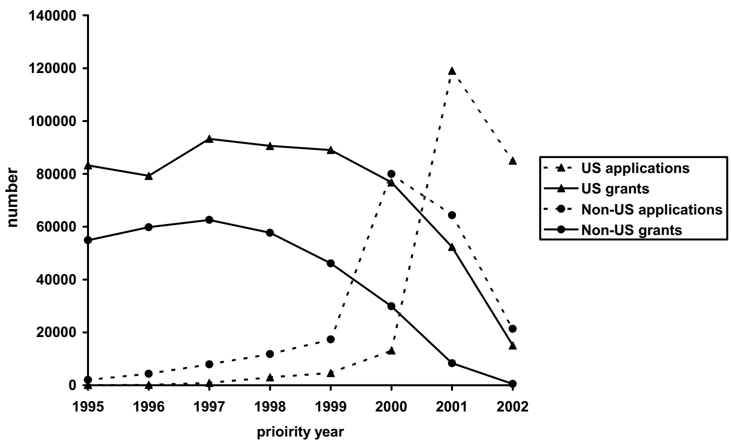


Figure 9.1. Patent applications and grants at the USPTO by US and non-US inventors by priority year.
Source: USPATFULL via Host STN

[3] See Section 2.4, for a short introduction to PCT applications.
[4] PCT applications are explained in more detail in Section 2.4.

2.2 References to Countries of Origin

Another factor influencing patent statistics is the assignment to countries of origin, which is certainly not a trivial task. Three different options are used in the literature[5].

First, a patent can be assigned to a country based on the office of the priority application, which is derived from the legal background. As described above, the priority filing is the first application of a specific patent; in many cases the first application is filed at the patent office of the country where the invention originated. This is the justification for using the priority country as the country of origin. The most important reason for this choice, however, is a practical one: all patent databases record the priority country, whereas other references are not always available.

Second, the inventor country information can be used for country assignment. Here the underlying argument is based on the concept of 'national systems of innovation' and the influence of their educational system, industrial relations, technical and scientific institutions, public policies, cultural traditions, etc. It is assumed that national systems of innovation are still the significant context for analysing technological developments (Grupp and Schmoch, 1999). Furthermore, it is also assumed that the location of the institution performing R&D and the inventor's place of residence, which together constitute the inventor country information, are generally identical. Inventor countries can be counted in different ways: a patent may be assigned to a country if at least one of the inventor residences is in this country. However, recently the share of patents with multi-national inventor teams has been increasing. In order to avoid double counting it is possible to refer to the first inventor only, or to use fractional inventor country counts. Both methods require the use of in-house databases. These more sophisticated approaches have visible effects on the absolute number of applications of specific countries, whereas the country relations are only slightly biased in favour of countries with a high degree of international cooperation. But this can be interpreted as an intended effect.

The third source of assignment is the applicant country. In this case the criterion of assignment is the ownership or the power of disposition. The country of the applicant is less clearly defined because in some international groups the parent companies apply for all patents, whereas in other ones the affiliations appear as the applicants.

[5] See also Dernis et al. (2001, p. 140).

Table 9.2. Number of applications at the EPO according to different methods of country assignment for selected countries, priority year 2000

Country	*Priority country*	*Inventor country*[6]	*Applicant country*
USA	32,441	29,765	29,804
Germany	19,617	22,536	21,293
Japan	19,210	20,533	20,401
Canada	360	1,793	1,468
Belgium	159	1,517	984

Source: PATDPA via host STN.

The effects resulting from using either of these three options on the outcome of patent analyses can be seen in Table 9.2. The most striking differences are between the results gained from an analysis based on the priority country versus the inventor country or applicant country data. Some countries, like Canada and Belgium, appear to be significantly under-represented if priority data are used as the basis. An additional, more detailed look at the data reveals that Belgian inventors tend to apply for their patents directly at the *European Patent Office* (EPO), thus getting a European priority, because the small domestic market is obviously less interesting for patent protection. For the priority year 2000, about 49 % of the patents filed by Belgian inventors were applied for at the EPO as the first patent office. As a result the priority country reference does not prove useful. A similar picture emerges for Canadian inventors. In this case the USPTO is by far the most important and thus the preferred patent office of first filing. In 2000 almost 74 % of Canadian inventors used the USPTO first to file their patent applications. First applications from Canada and some Asian countries explain the high number of US priorities compared to the US inventor and applicant data. As a general consequence it also proves problematic to use priority data as the source for country assignment. The relevance of the market appears to be an important factor for choosing the priority patent office, in particular where small countries are concerned, but larger countries are also influenced by this effect. In country comparisons small countries would be greatly underestimated, large countries moderately overestimated. A more consistent picture emerges when comparing data based on the inventor country and the applicant country. There are hardly any differences when comparing the figures for large countries, whilst again variances appear for small countries. To a large extent this difference can be explained by the effect of a dominance of large multi-national firms in smaller countries reflected in the applicant country. If the objective is to

[6] Designation by the criterion that at least one inventor resides in the country.

analyse the strength of a specific national R&D system, it is recommended to use the reference to inventor countries, as these come closest to assigning inventive activity to the location where it is carried out. Inaccuracies could appear with regard to regional analysis, e.g., when analysing patenting activities in the border region between Germany and Switzerland where a large number of employees of Swiss chemical enterprises live in Germany and thus appear with German inventor addresses (Hinze et al., 2002: 10f.).

If the first inventor or fractional inventor counts were used exclusively, the figures by inventor country in Table 9.2 would be slightly lower than the figures by applicant country for the larger countries USA, Germany, and Japan.

2.3 Offices of Reference

The selection of an appropriate patent office as a basis for patent statistics is another relevant factor which can significantly influence the outcome. Legal, cultural, and economic factors account for the differences between the various national patent systems.

Patent protection is always limited to the country where a patent is filed. The decision about the territorial coverage of a patent is an active one made by the applicant, and patents have to be applied for at the patent offices of the countries considered to be relevant. The selection always depends on the interest of the company in the respective market. If a patent is only filed in the home country, this decision reflects that only the domestic market is of interest for that particular product or process. If a patent is applied for in various countries it can be assumed that the company also intends to produce or distribute the product on these foreign markets. The applicant has to make a decision about applications in foreign markets within the first year after the first application, the priority year in legal terms. Then the priority date is taken as the novelty reference according to international rules. The priority year often proves to be quite short to decide on the international market–relevance of a specific product.

Often firms tend to apply to their national patent office first[7]. This gives rise to the problem that, if using data from national patent offices for statistical analyses, the technological strength of the home nation is generally overestimated compared to other countries owing to the so-called 'home advantage' effect (Schmoch et al., 1988, pp. 54 ff.). This effect becomes evident when looking at the data given in Table 9.3.

[7] This is, however, not always the case, as discussed in Section 2.1.

Table 9.3. Share of patent applications at various patent offices by country of origin, application year 2001.

Country of origin	*USPTO*	*JPO*	*DPMA*	*EPO*
USA	55.3 %	5.3 %	22.1 %	27.7 %
Japan	18.2 %	88.1 %	13.0 %	18.0 %
Germany	5.7 %	1.2 %	34.8 %	19.4 %
France	2.1 %	0.63 %	4.8 %	6.2 %
Netherlands	0.8 %	0.2 %	3.5 %	4.9 %
Switzerland	0.7 %	0.2 %	3.6 %	3.5 %
United Kingdom	2.5 %	0.5 %	3.2 %	4.4 %

Source: EPO, 2001; USPTO, 2001 and 2002; JPO, 2002; DPMA 2002.
USPTO = United States Patent and Trademark Office; JPO = Japanese Patent Office; DPMA = Deutsches Patent- und Markenamt; EPO = European Patent Office.

In each case, the respective home nation has by far the biggest share of patent applications at its own national patent office which underlines the importance of the respective domestic market. For instance, US-American applicants are responsible for more than 50% of all patent applications at the USPTO. Here European and Japanese players are consequently less prominent. Even more striking is the dominance of Japanese applicants at the Japanese patent office (JPO) with 88% of all patent applications.

Taking the JPO as an example, cultural and legal differences between the various patent offices also become clearly visible. The relevance assigned to patent applications in Japanese firms and the related incentive and reward system for employees in conjunction with the Japanese patent law yield particularly high numbers of domestic patent applications at the JPO. For instance, in 2001 at the JPO Japanese applicants filed almost 387,000 patents, whilst in the same period only 52,650 patents were applied for by German applicants at the German patent office (DPMA), even though the volume of industrial R&D in Japan is only about twice as high as in Germany.

However, most of the Japanese patent applications at the JPO are of comparably low value, and thus remain domestic patent applications only.

The choice of the patent office used for analyses is also highly relevant when comparing non-domestic countries. The US market, which is the single largest market in the world, is, of course, of special interest to firms, in particular to Japanese ones. However, the European market seems to be of comparable relevance because the share of patent applications from Japan at the USPTO and the EPO is almost identical (Table 9.3). However, the share of patent documents at a certain patent office is not sufficient to assess its importance. It is better to take the proportion of foreign applicants into account since shares are strongly influenced by the contribution of domestic applicants (see Table 9.4). From the data we can see that Japanese applicants at the USPTO are almost twice as numerous compared to the British activities at the EPO (see also below). For US firms the European market is very important. This is underlined by the USA being the country applying for the biggest share of patents at the EPO, amounting to almost 28%. Germany is the second largest country at the EPO with about 19%. European firms generally focus on the European market. Their technological strength would be underestimated if only data from the USPTO were considered. For instance, Swiss firms contribute less than one per cent of the patent applications at the USPTO, but their share at the EPO amounts to 3.5%. Another example is the relation of patent shares between France and the United Kingdom. According to the data for British firms, the number of patents at the USPTO is slightly higher than for French firms, whilst the opposite is true at the JPO. Thus the data indicate that the Japanese market seems to be more important for French firms than it is for British firms, whereas the picture is reversed on the US market. Even more obvious is the stronger orientation of French firms towards the European market; whilst French players contribute 6.2% of the patent applications at the EPO, the share of British filings is only 4.4%.

The variations owed to the selection of the patent office when assessing the technological strength of individual countries are illustrated from a different perspective by the data in Table 9.4. Here we used the number of British patent applications at various patent offices as a reference; the strength of other countries was calculated in relation to the British activities. It can be shown that the results differ depending on which reference office is selected. The home country is clearly overestimated, as has already been mentioned above.

Table 9.4. Patent activity of selected countries in relation to British patents at various patent offices, application year 2001

Country	*USPTO*	*JPO*	*DPMA*	*EPO*
USA	22.5	10.7	6.8	6.3
Japan	7.4	176.9	4.0	4.1
Germany	2.3	2.5	10.8	4.4
United Kingdom	1.0	1.0	1.0	1.0
France	0.8	1.3	1.5	1.4

Source: EPO, 2001; USPTO, 2001 and 2002; JPO, 2002; DPMA 2002; own calculation. Abbreviations as in Table 9.3.

Amongst the potential offices of reference the EPO appears to be the most balanced one with regard to country comparisons. The distortions here are the lowest because the EPO is not a national but a regional/international patent office covering the European market. The EPO provides a standardised central grant procedure for all member countries of the European Patent Convention (EPC) selected as designation countries. Europe is an interesting and important market for North-American and Japanese firms and US-American and Japanese players contribute significantly to the patent applications at the EPO. However, even though the bias produced by the 'home–advantage' is reduced, the activities of European countries, in particular of Germany, are still over-represented at the EPO. This bias may be called 'regional advantage', but it is substantially lower than the 'domestic advantage'. According to our experience the regional advantage is even lower in research–intensive areas, as overseas countries, such as the United States, tend to apply for a higher share of their domestic applications in Europe as well. The balanced structures at the EPO also support the calculation of specialisation indices as suggested by Soete and Wyatt (1983), because the average at the EPO is not dominated by any one country. Furthermore, the distribution of patent applications by country is quite similar to the distribution in terms of foreign trade or R&D expenditure (Schmoch and Frietsch, 2001).

2.4 Reference to Multiple Offices

The domestic advantage at national patent offices proves to be a major problem when comparing the performances of different countries. This is why various authors suggested quite early that foreign patents should be used as a reference (Basberg, 1983; Schmoch et al., 1988, pp. 54 ff). However, the problem remains that at a specific office all the countries can be compared except the home country concerned. Various approaches with reference to international patenting have been conceived to cope with this shortcoming.

Gerstenberger (1992) used all inventions which had patent applications in at least two countries in order to achieve comparability between the countries of origin. At first sight this criterion ensures equal conditions for all countries and is convincing because of its simplicity. However, if the consequences of its application to specific countries are considered in more detail, the underlying inconsistencies become clear. From the perspective of a European country the first foreign application is generally registered in another European country. In the case of Japanese firms the first foreign applications are filed to the United States, a decision which is more difficult to take than in the case of a European firm with regard to a geographically neighbouring country. Furthermore, the approach mixes patent applications in Europe and Japan with patent grants at the USPTO, different types of patent documents, so that the outcome can hardly be interpreted in an appropriate way.

Grupp et al. (1996, pp. 279 ff) proposed the use of so-called triad patents to overcome the drawbacks of Gerstenberger's approach. Triad patents refer to patents which are applied for at all major offices of the triad regions, the USPTO, the JPO, and the EPO. This concept specifies the geographic location of foreign patents more precisely and the patent indicators have a closer link with international competitiveness in terms of foreign trade with technology–intensive goods. But again, the inclusion of the USPTO in the triad concept appears to be a major problem because it involves a mixture of grants and applications and a publication delay (Dernis, 2003). However, it is possible to circumvent these problems by looking more closely at international patent flows; the reference data shown in Figure 9.2 record applications, not grants, for the three triad offices. If a European applicant decides to patent in overseas countries, the first choice is the United States. Only about 58 per cent of the applications in the United States are registered in Japan as well. From the Japanese perspective the United States are again more important than Europe with regard to foreign patents. Only about one third of the applications registered at the USPTO are also filed at the EPO. Thus in both cases the applications at the USPTO are not the restrictive factor in the triad analysis. So it is possible, in fact, to limit the 'triad' searches to the JPO and the EPO because it can be reasonably assumed that all foreign applications in Japan and Europe are also registered in the United States. In this way the mixture of the different types of patent documents and the publication delay can be avoided.

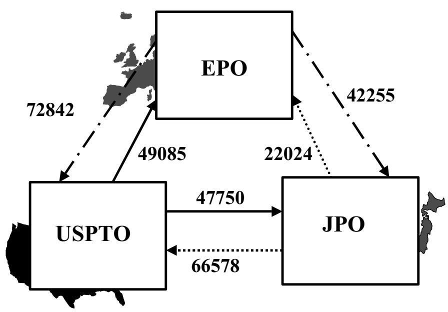


Figure 9.2. Flows of patent applications from Japan, the United States and the EPC member countries to the major triad patent offices, application year 2000.
Source: PATDPA via Host STN, EPO et al. (2001)

Even in the 'reduced' version, the practical application of the triad concept requires a sophisticated database which makes it possible to link Japanese and European patents. In such a database so-called patent families have to be recorded in a consistent way[8]. Only a few, quite expensive databases permit such family searches.

Against this background, analyses at the EPO are still an attractive alternative, as explained in section 2.3. Recently, international patent applications under the so-called Patent Co-operation Treaty (PCT) have become increasingly important. The PCT process is linked to the World Intellectual Property Organization (WIPO) in Geneva, Switzerland, but the applications are made at already existing patent offices such as the USPTO, JPO, EPO etc. Although the PCT process had already been introduced in 1978 as a main international patent application system[9], it only achieved a broader acceptance among applicants in the middle of the nineties. Since then the number of PCT applications has steadily increased owing to their advantages, such as the reduction of cost risks linked to foreign patenting (for more details, see Schmoch, 1999b). For a long time PCT applications could not be used for statistical country comparisons, because some relevant countries such as Japan, Switzerland, Italy, or South Korea only had low shares of PCT applications within their foreign patenting. Now every country has since adopted the PCT system, so that the relations between countries are more realistic. Because of the long transitional phase involved, PCT applications cannot be used for analyzing time series, so that a pragmatic solution would be the combination of PCT data to determine country relations and EPO data to analyse time trends. In any case, the analysis of PCT data at the recent edge is an interesting new possibility for country comparisons which may be able to replace more complex approaches such as the 'at least one foreign country' or the 'triad' concepts.

3. INDICATORS OF QUALITY

In this section we will present several indicators of the quality of patents. We do not aim to introduce novel indicators, but rather to discuss the appropriateness of indicators already established in the literature with regard to country comparisons and practical purposes.

A standard indicator of quality is the grant as opposed to a pure application (see, Guellec et al., 2000). Although there is some evidence that granted patents have a higher technological value than applications, practical

[8] The original (domestic) application and all foreign patents referring to this constitute such a patent family.

[9] It is not possible to grant patents within the PCT process. All PCT applications have to be transferred to national or regional offices for the further grant procedure.

considerations strongly support the use of applications. First, patent applications are examined with regard to their technical novelty and the inventive step, not with regard to their economic value. Second, it is rather difficult to interpret differences in grant shares in an appropriate way. Third, and this is the most important aspect, the time lag between the publication of an application and its grant is often considerable, as discussed above (Section 2.1.).

The effect of application versus grant counting is illustrated in Figure 9.3. We chose the year 1994 in order to obtain a large enough time lag to the date of search and thus a sufficiently high grant share. This was 63 per cent on average, but in our experience this will rise to about 70 per cent if the time lag is increased by another five years. Most of the countries considered have grant rates close to the average. Major exceptions are Germany and France with 72 and 70 per cent, respectively, versus 51 per cent of the United States. So the introduction of grants actually discriminates between countries. But it is not proved that a higher grant rate is linked to higher quality. The described effect may be owing to European applicants' better knowledge of the examination requirements at the EPO, whereas US inventors have to transfer US application documents into a European structure. In addition, the increasing use of PCT applications and the referring delayed transfer to the EPO may play a role. Against this background we recommend the analysis of patent applications, not grants, in order to be as topical as possible.

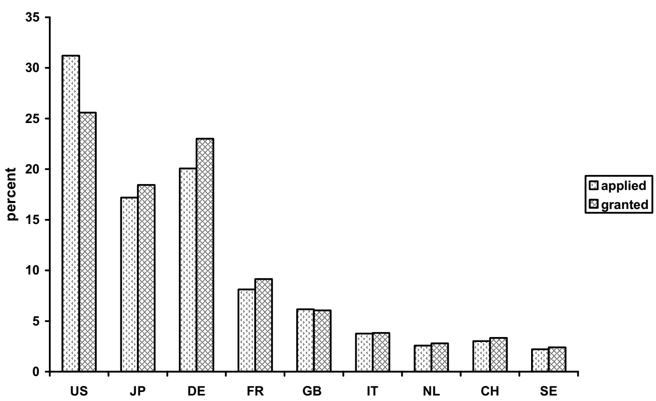


Figure 9.3. Share of selected countries within all applications or grants at the European Patent Office, priority year 1994 (Update: Dec. 2003)
Source: PASTDPA (Host STN), own computation

Many authors suggest using citations in subsequent patents as a quality indicator (for instance Narin, 1987; Carpenter and Narin, 1983; or

Trajtenberg, 1990). This approach discriminates between countries much better than the grant criterion. The major shortcoming of this method is again the delay between the publication of the citing and cited documents, so that the value of patents can only be assessed for less recent years. However, highly cited patents do seem to have a long-term impact on the performance of firms (see the contribution of F. Narin in this book), so that this indicator still proves useful.

Harhoff et al. (2003) found a significant correlation between the value of patents and the number of citations in their search reports (backward citations). If this result is confirmed by other studies this indicator would be an interesting alternative for (forward) citations by subsequent patents, because the delay between application and the first search report is much shorter than in the case of forward citations.

A further frequently used quality indicator are foreign patent applications (Soete and Wyatt, 1983) which are published strictly 18 months after the priority application, generally the first domestic application. A high quality of foreign applications can be assumed, because these involve substantial additional costs compared to purely domestic patents (see, Archibugi, 1992).

As explained in sections 2.3 and 2.4, it is not reasonable to use the simple fact of a foreign application for country comparison. Rather, we suggest referring to a specific patent office, where the frame conditions are well known, and the geostrategic choices of applicants from different countries can be taken into account in a rational way.

An extension of the concept of foreign application is the size of families (Faust 1990), the number of foreign applications at different offices linked to a specific invention. Harhoff et al. (2003) found a significant correlation between family size and value of patents, a result which supports the use of this approach. However, geo-strategic considerations are significant here as well, and in consequence, the family size is useful for discriminating within countries by quality, but not for doing so between countries.

Similar considerations apply to the size of inventor teams as a quality indicator (Schmoch et al., 1988, p. 63). The organization of research teams and the designation of inventors depend heavily on national (cultural) frame conditions, and country comparisons prove to be less meaningful.

A further quality indicator recently suggested by Harhoff et al. (2003) is opposition in the context of the patent grant procedure at the EPO, which displays a highly significant correlation with the (economic) value of patents. This indicator is available at an earlier point in time than (forward) citation rates by subsequent patents and is therefore of special interest. Recent discussions with industry experts seem to indicate a change of behaviour with regard to opposition, at least in some industrial sectors. Instead of diminishing the scope of a competitor's patent or even destroying it, some firms seem to prefer to avoid opposition and instead cross-license with their competitors. It remains to be seen whether this change of attitude has any significant impact on the statistical relation between opposition and value.

It is, of course, desirable to have quality indicators which are available as early as possible. But we have to ask at which point in the life cycle of an invention can its quality be determined in a reliable way? In the case of patents quality primarily refers to their technical and, in particular, their economic value. Whether an invention is successful in the market place depends on its technical quality, which may be reflected in the patent document, and on its economic success, which is largely a reflection of consumer decisions. The market penetration of an invention generally only becomes visible a considerable time after the first patent application. To a large extent, the market value is already covered by 'late' quality indicators such as (forward) citations. 'Early' quality indicators may express a higher potential of future market success, but their prognostic power will always be limited in principle.

4. SEARCH STRATEGIES FOR STATISTICAL ANALYSES

Statistical patent analyses often refer to specific technological areas. In many studies these areas are labelled using a general term such as 'biotechnology', 'information technology', etc., without a detailed description of the underlying definition. However, an adequate definition is crucial. For this purpose some basic rules of sample definition have to be followed.

In database searches it is nearly impossible to cover all relevant patents[10], because they can appear in unforeseen contexts. Therefore the aim of a patent search is to define samples which are statistically representative for the domain considered. Therefore the samples should be as large as possible and include as few inappropriate documents as possible. This orientation implies a major difference between statistical searches and novelty searches for legal purposes, in the examination process of a patent application. The aim of legal searches is to identify all the relevant documents, even if they appear in unforeseen, marginal contexts. Against this background, legal searches often include unsuitable documents, which are then excluded by a process of intellectual examination of each document covered by the original sample. This can be realised in legal searches, because the search areas are

[10] The term 'patent' includes 'patent applications' and is used to simplify the text.

more precisely defined and much smaller than in statistical searches. In statistical analyses the number of documents is much larger, so they have to rely primarily on automatic searches without detailed intellectual or manual examination.

The principles of statistical searches are illustrated in Figure 9.4 in which the light grey area represents the complete document set of the technical area considered. The standard approach in patent searches is to look at appropriate codes of the International Patent Classification (IPC), with which the examiners of the patent office classify each patent application. Each patent document has at least one IPC code, the primary code, and often secondary codes too. Where suitable IPC codes exist it is strongly recommended to use primary and secondary codes in order to achieve an optimal yield.

In order to identify appropriate IPC codes for defining a specific area, it is necessary to scan the whole IPC classification, which is hierarchically built, in a top-down approach. The identification of appropriate codes requires, of course, sufficient technical knowledge of the examined area. It is helpful to start the search with adequate keywords and to look, by statistical analysis, at the IPC classes or sub-classes in which these keywords are used most frequently. This approach is a practical support tool, but again, the outcome has to be assessed intellectually before the identified IPC codes can be integrated into a search strategy.

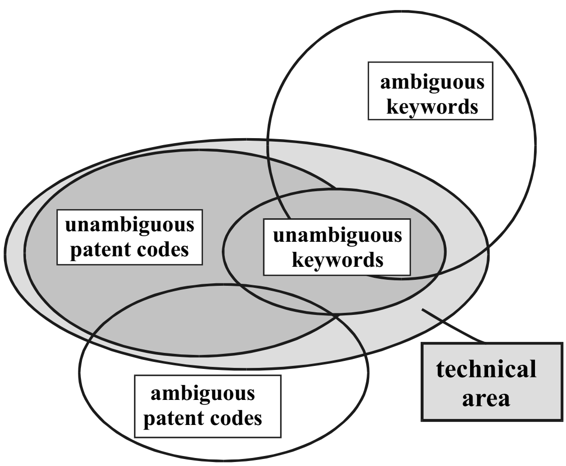


Figure 9.4. Illustration of data sets identified in statistical patent searches

In general the share of the technical area covered by IPC codes is very high, although some relevant documents may not be included. To find additional relevant documents keyword searches can be performed. It is important that the keywords used in such additional searches are unambiguous with reference to the area considered, so that exclusively relevant documents are selected. Experts in a specific field tend to cite keywords which are often used in their field, but also in other contexts. These keywords would result in too many unsuitable documents, so that the sample would not be representative for the field considered. This situation is illustrated in Figure 9.4 by the area of ambiguous keywords. The same applies to ambiguous classes covering relevant documents to a certain extent, but also documents of other areas.

In some cases, however, it can be useful to include ambiguous keywords or IPC codes if the share of unsuitable documents is limited and does not call the statistical representativeness of the total sample into question. Furthermore, the precision of ambiguous keywords and IPC codes can be improved by combining them with other keywords or IPC codes. This is done by including positive associations (by intersection) or excluding negative associations. In Figure 9.4 the area shaded dark grey represents the final patent sample consisting of documents with unambiguous patent codes and keywords[11]. This area covers the largest part of the total set of relevant documents, shown in light grey.

Keyword searches in official patent documents published by national or regional patent offices are generally less productive, because the legal requirements of disclosure with regard to titles and abstracts are not very strict. If a search has to be based on keywords to a considerable extent, it should be executed using the so-called World Patents Index (WPI). The WPI records contain improved titles and abstracts describing the technological content of each patent application.

To sum up, patent searches are generally based on IPC codes complemented by keywords. Statistical searches aim at the definition of representative samples and cannot cover all the relevant documents in most cases. It is essential to avoid ambiguous IPC codes and keywords, as a high share of unsuitable, misleading documents can bias the outcome. Finally, it is important to look at the use of appropriate databases in the case of keyword searches.

5. CONCLUSIONS

The outcome of statistical patent analyses largely depends on the choice of appropriate reference. We suggest the priority year as the time reference

[11] Keyword and patent code combinations are not shown in order to simplify the picture.

because it is closely related to the time of invention. For defining the country of origin the best reference proves to be the inventor country, which reflects where the invention was made. As to the office of reference, it is important to be aware of domestic or regional advantages and of the geo-strategic position of the office. For balanced country comparisons the *European Patent Office* appears to be suitable, but the so-called triad concept is more closely linked to the distribution of R&D activities. A new alternative is presented by the growing relevance of international PCT applications, the analysis of which may replace more complex approaches.

With regard to the quality of patents, it has to be taken into account that the indicators reflect technological as well as economic aspects. The technological quality may be assessed by early indicators examining the content of the application documents. The economic quality is determined by the market success, which is very difficult to foresee. Late indicators such as citations or opposition already reflect the aspect of market success.

The appropriate definition of samples largely determines the outcome of statistical patent analyses. It is vital to work with statistically representative samples. Simply trying to obtain as large a sample as possible carries the inherent danger that the proportion of inaccurate documents will be too high and will therefore result in misleading figures. A suitable definition requires sufficient expertise in patent classifications and good technical knowledge.

All in all, it is important to document carefully the methodological choices of a statistical patent analysis and to discuss the implications for the outcome. Only on this basis is an appropriate interpretation of the results possible.

ACKNOWLEDGEMENTS

The article is based on experience in several projects on behalf of the German Ministry for Education and Research (BMBF), the European Commission (DG Research), and the OECD.

REFERENCES

Archibugi, D. (1992). Patenting as an indicator of technological innovation: A Review. *Science and Public Policy*, 10, 357–368.

Basberg, B.L. (1983). Foreign patenting in the US as a technology indicator. *Research Policy*, 12, 227–237.

Basberg, B.L. (1987). Patents and the measurement of technological change: a survey of the literature. *Research Policy*, 16, 131–141.

Campbell, R.S., Nieves, A.L. (1979). *Technology indicators base don patent data: The case of catalytic converters*. Richland: Batelle Pacific Northwest.

Carpenter, M.P., Narin, F. (1983). Validation study: Patent citations as indicators of science and foreign dependence. *World Patent Information*, 5, 180–185.

Department of Commerce (2000). Changes to implement eighteen-month publication of patent applications. *Federal Register*, 65 (183). September 20, 2000, 57024-57061. Washington, D. C.: DoC.

Dernis, H. (2003). *OECD triadic patent families*. Presentation at the WIPO-OECD Workshop on 'Statistics in the Patent field' in Geneva. September 18 and 19, 2003.

Dernis, H., Guellec, D., van Pottelsberghe, B. (2001). Using patent counts for cross-country comparison of technology output. *STI Review* (OECD), 27, 129–146.

DPMA (2002). *Annual Report 2002*. München: DPMA.

EPO (2001). *Annual Report 2001*. Munich: EPO.

EPO (2002). *Annual Report* 2002. Munich: EPO.

EPO, JPO, USPTO (2001). *Trilateral Statistical Report.* 2001 Edition. Munich, Tokyo, Washington, D.C.: EPO, JPO, USPTO.

Faust, K. (1990). Early identification of technological advances on the basis of patent data. *Scientometrics*, 19, 473–480.

Gerstenberger, W. (1992). Zur Wettbewerbsposition der deutschen Industrie im High-Tech-Bereich. *Ifo Schnelldienst*, 13, 14–23.

Griliches, Z. (1990). Patent statistics as economic indicators: A Survey. *Journal of Economic Literature*, 28, 1661–1707.

Grupp, H., Schmoch, U. (1999). Patent statistics in the age of globalisation: new legal procedures, new analytical methods, new economic interpretation. *Research Policy*, 28, 377–396.

Grupp, H., Münt, G., Schmoch, U. (1996). *Assessing different types of patent data for describing high-technology performance*. In OECD (Ed.), Innovation, Patents and Technology Strategies (pp. 271–284). Paris: OECD.

Guellec, D., van Pottelsberghe de la Potterie, B. (2000). Applications, grants and the value of patents. *Economic Letters*, 69, 109–114.

Harhoff, D. Scherer, F. M., Vopel, K. (2003). Citations, family size, opposition and the value of patent rights. *Research Policy*, 32, 1343–1363.

Hinze, S., Jappe, A., Koschatzky, K. (2002). *International Benchmark Club. Innovationsstatistik. Vorstudie Patente und Bibliometrie*. Karlsruhe: Fraunhofer-ISI

JPO (2002). *Annual Report 2002*. Part III Statistics. Tokyo: JPO.

Kuntze, U., Müller, P., Pfeiffer, R., Rempp, H., Westermann, G. (1975). *Erfassung und Bewertung technologischer Entwicklungstrends auf der Basis des im Deutschen Patentamts vorhandenen Datenmaterials und technologischen Wissens*. Karlsruhe: Fraunhofer-ISI.

Narin, F., Noma, E., Perry, R. (1987). Patents as indicators of corporate technology strength. *Research Policy*, 16, 143–155.

OECD (1994). *The measurement of scientific and technological activities. Using patent data as science and technology indicators*. Patent Manual 1994. Paris.

Pavitt, K. (1985). Patent statistics as indicators of innovative activities: possibilities and problems. *Scientometrics*, 7, 77–99.

Schmoch, U. (1999a). *Eignen sich Patente als Innovationsindikatoren?* In R. Boch (Ed.), Patentschutz und Innovation in Geschichte und Gegenwart (pp. 113–126). Frankfurt a. M.: Verlag Peter Lang.

Schmoch, U. (1999b). Impact of international patent applications on patent indicators. *Research Evaluation*, 8, 119–131.

Schmoch, U., Frietsch, R. (2001). *Patentaktivitäten: Dynamik und Spezialisierung im internationalen Längs- und Querschnitt*. In NIW, DIW, Fraunhofer-ISI, ZEW, Stifterverband (Ed.), Indikatorenbericht zur technologischen Leistungsfähigkeit Deutschlands 2000/2001 (pp. 100-112). Hannover: NIW.

Schmoch, U., Grupp, H., Mannsbart, W., Schwitalla, B. (1988). *Technikprognosen mit Patentindikatoren*. Zur Einschätzung zukünftiger industrieller Trends bei Industrierobotern, Lasern, Solargeneratoren und immobilisierten Ernzymen. Köln: Verlag TÜV Rheinland.

Schmookler, J. (1966). *Invention and Economic Growth*. Cambridge (Mass.): Harvard UP.

Soete, L.G., Wyatt. S.M.E. (1983). The use of foreign patenting as an international comparable science and technology output indicator. *Scientometrics*, 5, 31–54.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of inventions. *RAND Journal of Economics*, 21, 172–187.

USPTO (2001). *Performance and Accountability Report: Fiscal Year 2001*. Washington, D.C.: USPTO.

USPTO (2002). *Performance and Accountability Report: Fiscal Year 2002*. Washington, D.C.: USPTO.

Chapter 10

SCIENCE MAPS WITHIN A SCIENCE POLICY CONTEXT

Improving the Utility of Science and Domain Maps Within a Science Policy and Research Management Context

Ed C.M. Noyons

Centre for Science and Technology Studies (CWTS), Leiden University, the Netherlands
E-mail: noyons@cwts.leidenuniv.nl

Abstract: Science mapping within the science policy and research management context has had a promising start during the seventies, but it lasted until the early nineties before useful technology was developed to make it really work. Recently domain visualization seems to have become a mature research area. Therefore, the actual potential of visualization for this particular purpose can be explored in more detail. In this chapter a short history is outlined and the current possibilities and requirements of contributing to a sound evaluation tool in the near future are listed. By using a case study the potential of bibliometric maps is illustrated within a policy context. It shows how these maps may be used to visualize the research focus of actors within one field and to compare them. The main issue raised in this chapter is the requirement of reference points in order to make useful comparisons.

1. INTRODUCTION

Maps of science are an appealing representation of academic knowledge and as such applicable in many contexts. In this chapter I discuss in short the history of science mapping and, in particular, within the science policy and research management context. An obvious link is made to *domain visualization* techniques that have been developed in the past decade.

Since the introduction and hectic starting phase of bibliometric maps as representations of science fields and research areas, there has been a long time with hardly any developments. Only since the availability of graphical user interface techniques in the nineties, has this kind of maps been

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 237-255.

developed further. The appealing aspect of using the map as an interface to find information or investigate the generated structures has contributed to this revival considerably.

It is now time to define for each application area the proper requirements for these interfaces in order to be able to stay on the right track. There is still a lot misunderstanding about the validity and utility of the maps (c.f., Noyons, 1999 and 2001), but with clear guidelines we should be able to make a major progress in the next ten years. In this chapter the focus is on the specific applications of these maps within a science policy and research management context.

2. MAPPING

A science map is two- or three-dimensional representation of a science field, a 'landscape of science', in which the items in the map refer to themes and topics in the mapped field, such as cities on a geographical map. In these maps the items are positioned in relation to each other in such a way that the ones which are cognitively related to each other are positioned in each other's vicinity, whilst the ones that are not or hardly related are distant from each other.

The maps of science considered in this chapter are those based on bibliographical data, the bibliometric maps of science. As scientific literature is assumed to represent scientific activity (Merton, 1942; Ziman, 1984), a map based on scientific publication data within a science field *A* should be able to represent the structure of *A*. It will depend on the information (elements of a bibliographical record) used to construct the map, what kind of structure is generated, where kind of maps refers to aspects like cognitive and social structures.

Most science maps are constructed by the co-occurrence information principle, i.e., the more two elements occur together in one and the same document the sooner they will be identified as being closely related. The science mapping principle dictates that the more related two elements are the closer to each other they are positioned in a map, and the other way around: the less related two elements are the more remote they will be in the map.

Different elements of a bibliographic record may be used to generate a structure. Each element reveals a specific aspect of a publication and can therefore be used to compile a specific kind of structure, unique in a sense, but through the coherence of a publication always related to the structures based on other elements.

A bibliographic record (representing the publication) contains a range of elements. The important ones are:

– Author(s) of a publication;
– Title of a publication;
– Source in which the document is published, e.g., the journal, proceedings or book;
– Year of publication;
– Address(es) of the (first) author(s);
– Abstract of the publication.

In specialized bibliographic databases other information may be included as well:

– Cited references;
– Publisher information of the source;
– Keywords provided by the author or journal editor;
– Classification codes added by the database producer;
– Indexed terms added by the database producer.

In principle all these elements can be used to build a map. As mentioned above, the kind of structure that is generated depends upon the element used. A map based on co-occurrence of *authors* is more likely to unravel a *social* structure (c.f., Peters and Van Raan, 1991) of a science field, describing the relations between people. A map based on co-occurrences of *classification codes* rather describes the relation between content elements and thus reveals a *cognitive* structure.

A map as such has the function of being an aid to help users to explore a publication collection (e.g., search results, a science field). It will depend on the kind of user and the objective as to which map fits best.

In general, all mapping and visualization techniques structure large amounts of data by clustering documents, in our case publications. They may be represented in more than one cluster if the clusters overlap. The clusters are defined by elements taken from the publication data (more accurately, bibliographic data), and depending on the objective of the map a choice for one particular element is made. To provide a better view of the possibilities, I compiled a simple model to represent scholarly communication, as far as relevant for scientometric studies from the point of view in this chapter.

The model shows three important elements involved in the entire process of communication, and relevant for scientometric and bibliometric research:

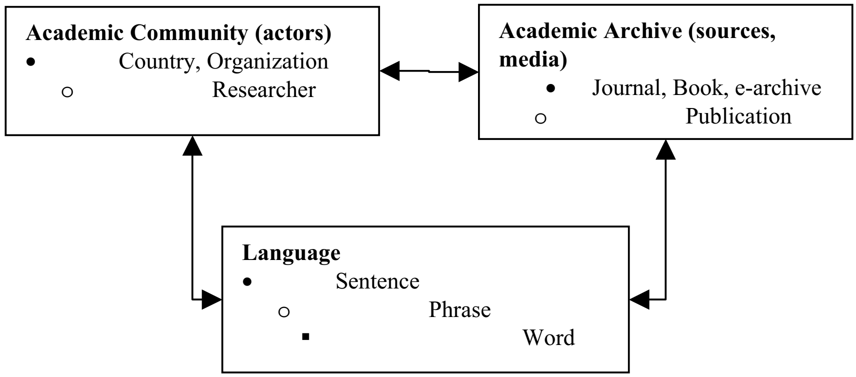


Figure 10.1. Scholarly communication model applicable to scientometrics.

Other elements are involved in the process, e.g., science policy, technology, and societal issues, but I do not take them into consideration here. For a much more complex and broad model in which this may fit I refer to Wouters (1998).

Each of these three elements shows a varying level of aggregation, down to the smallest element applicable to the present research of scientometrics, being the researcher, the publication, and the single word.

Most research performance evaluation studies in which the output and impact of actors (countries, organizations) are evaluated involve *archive* and *community*. Co-word analysis involves *language* and *archive*. Semantic analyses of a scientific author involve *language*, as taken from publications, and linking the other elements (*archive* and *community*). Studies of journals (e.g., journal impact factors) have (parts of) the *archive* as a starting point. Every journal, such as the entire set of publications represented by the *Science Citation Index*, represents, in fact, a sub-collection of the entire archive. Citation data is primarily restricted to the archive, in which one publication refers to an other.

Scientometric and bibliometric research apply to the *archive*, as publication collections are the starting point. Performance analysis concerns the evaluation of entities from the *community*, using the link between *author names* in publications and *researchers* or *affiliations* in publications and *organizations*. In co-citation analysis links within the archive are the basis, whereas author co-citation integrates community entities (*researchers* as *authors*).

Mapping techniques typically structure (parts of) the archive referring to entities from one of these three elements. Subsequently the structure can be linked back to either three, depending on the purpose of the map. For instance, in an actor analysis (at the level of country, organization or author) the activity or performance of a (list of) actor(s) is assessed. An actor (as a part of the academic community) is linked to the archive by looking at the author name or corporate source data in publications. By using the structure of the archive linked to the actor(s) we are able to characterize the actor's activity or performance. The structure we apply to the archive collection may be generated by using language elements, such as words or phrases, extracted from the archive (publications).

3. HISTORY

The history of mapping of science in the context of science policy and research management is relatively short. Moreover, it appears that only a few persons or groups have been engaged with this particular application of bibliometric maps. A highly interesting introduction to the history is found in Small (2003). Small sketches the birth of co-citation within a philosophical context of research paradigms, where the highly cited papers are concept symbols and clusters of these symbols should represent paradigms. With this method it should be possible to map the entire archive of science. The 'war maps of science' to be applied by 'government bureaucrats' were never really established, but the idea of doing so puts it in this particular context of science policy and research management. The development of co-citation analysis was adopted in many studies but the fundaments were only further developed at ISI by Henry Small (Small, 2003).

A spin-off of the co-citation approach is the analysis based on cited author names. This method was developed by White and Griffith (1981) to map the intellectual base of science fields (sub-collections of the archive). This technique uses the co-occurrence of author names in the reference list of articles. The more often two authors are co-cited the closer their intellectual relation. This should reveal the intellectual structure of a field. White and McCain (1998) developed this technique further and Author Co-citation Analysis (ACA) became a very popular technique for mapping an intellectual structure (amongst many others: Persson, 1994 and Persson, 2001).

In the early eighties a mapping methodology for a similar objective was developed by the École des Mines in Paris as described in Callon et al. (1983) and Callon, Law, and Rip (1986). They applied co-word analysis rather than co-citation, identifying clusters of words representing research themes. Callon mistrusted the citing behaviour of scientists and therefore considered words as a more appropriate element for creating structures of science. The well-known *Leximappe* software has been used in many studies since then but was not further developed after the early nineties. Based on the work of this French group, Kostoff created his Database Tomography (DT) and applied it in many studies. Unlike most other co-word studies, Kostoff incorporates no visualization techniques.

During the eighties and nineties most developments of the techniques mentioned were put on hold and it was not until the spread of graphical interfaces, such as HTML, Java applets and other Web interface techniques, that the mapping experienced a revival. This particular technology made it possible to use the maps as an interface and to explore the results and underlying data, rather than to look at hard copy structures only. In Lin, White, and Buzydlowski (2003) even a real-time version of the ACA application is presented. But there have been many more similar developments recently (see next section). The graphical user interface has opened doors for the Information Retrievalists and created new opportunities. In particular, the Self-Organizing Maps of Kohonen (1990) are adopted in many visualization tools (e.g., Polanco, François and Lamirel, 2001). Together with the revival of domain visualization, fundamental issues are raised (e.g., Cahlik, 2000; Ahlgren, Jarneving, and Rousseau, 2003; White, 2003).

Despite the revival of visualization techniques in the recent past, an increased application of the resulting maps within a science policy context is still not detectable. It seems that the application of the interfaces is most prominently focused on Information Retrieval; however, often inadvertently. The fundamental difference between these application areas is that an IR user uses the domain map to find an area of interest and zoom-in on a particular sub-area, down to the individual publications. A policy–related user may zoom in into sub-areas but will normally return to the overall view to draw conclusions.

4. CURRENT TECHNIQUES

Börner, Chen and Boyack (2003) — hereafter BCB — provide an excellent overview of the techniques used presently to visualize domains as a follow up of White and McCain (1997). All the techniques discussed could be used as an aid to search for relevant data. But it should be mentioned that most applications are developed to be used in IR, rather than in a science policy context. All discussed techniques are designed to create a domain map, i.e., a two- or three-dimensional representation of information entities to represent the structure of a domain. A domain can be a research field, or any other bibliography, or collection of publication data.

Most domain visualization studies seem to follow a similar path to go from publication data to the actual visualization, the map. In this process flow BCB discern the following steps:

1. Data extraction;
2. Unit of analysis (choice for author, term, document to be used to create the map);
3. Measures (statistics, unit of aggregation, thresholds to decide which items to use, etc.);
4. Similarity method (co-occurrence or other relation, matrix type, correlation measure);
5. Ordination (dimensionality reduction technique, clustering, scaling technique);
6. Display.

Different techniques may use different options at the described stages. And apart from the fundamental discussion about certain steps in a methodology (e.g., Ahlgren, Jarneving, and Rousseau, 2003; White, 2003), the choices to be made at several stages will also depend on the objective of the tool. Author names, in a co-author analysis, reveal a social structure, because they represent persons. This social structure may coincide with a cognitive structure, but that is another issue. Moreover, the interface for displaying the results should be fit to provide relevant answers to the questions at stake. This means that the different present techniques described by BCB may all have their own application area. In most cases they were primarily developed for Information Retrieval. They should be able to visualize a data collection in such a way that the user can zoom in on an appropriate sub-collection. This is an important issue because it means that this application area explicitly requires the entities behind the structure to be individual publications. In some cases the entities on the map are individual publications, in others the entities directly refer to a limited sub-collection of individual publications (co-citation map). But in many cases the entities in the map represent sub-collections of publications. Author co-citation clusters represent a structure of the publication oeuvres of co-cited authors. The research front of a co-citation map represents a structure of the collection of articles citing particular co-citation clusters. Similarly, in co-word maps words, terms, or clusters of words represent publications containing them. The words or terms in the maps do not mean anything as such. In terms of the model in Figure 10.1 an entity in a map always represents a sub-collection of the archive. The labelling of the entities is in most cases a language element or an entity or sub-collection of the community.

BCB explicitly mentions two approaches as being used in a science policy and research management context (Noyons, Moed, and Van Raan, 1999; and Boyack, Wylie, and Davidson, 2002). Still, I think that also the studies of Kostoff (e.g., Kostoff, Eberhart, and Toothman, 1998), Bhattacharya and Basu (1998), Widhalm et al. (2001), Schwechheimer and Winterhager (2000) and Salvador and Lopez-Martinez (2000), to mention a few relatively recent studies, are typically designed to be used in that context. And often they mention this orientation too. It seems, however, that in these cases the analysts did not always adjust the display (the final stage of the process flow) to disclose this option explicitly.

In the BCB review the results of many different techniques (both co-citation and co-word–based) are compared with each other applied to a collection of publications in the field of scientometrics and domains visualization. This collection, named ARIST (after the serial in which the review appeared) is analysed with all techniques available to them (either by software package or expertise) and results are compared. They conclude that every technique has its own special features, advantages, and disadvantages, and that it depends heavily on the objective of which one to use. Moreover they advise exploring results from more than one technique in order to provide insight from different perspectives.

For their study they only included techniques which are typically designed to be used for IR. And although there are obvious possibilities, the techniques are not designed to be used in a science policy context, neither are they reviewed within that context.

5. MAPPING AS A SCIENCE POLICY TOOL

As outlined above, a domain map is often used in contexts other than in the context of science policy/research management. It seems, however, that many issues brought up in domain mapping studies relate to policy–relevant questions. Especially the issues concerning the dynamics of a field under study and the issues concerning the actors (countries, organizations, or authors) relate closely to science policy and research management. Unfortunately, the applied techniques for visualising the issues do not always provide the proper maps or displays to contribute sufficiently to the discussion. In many cases there is a possibility of viewing time series of domains and hence to describe in detail the changes over time, but this approach lacks the means of a comprehensive overview of the relevant changes. These are not the Price's war maps of science on the basis of which science policy can prepare a strategy.

In the process flow of domain visualization (see previous section) there are two stages which seem most important to validate the use of domain maps in a science policy context: the choice for a unit of analysis and the stage of display[1] (more accurately, the design of an interface to use the map). As mentioned before, all techniques cluster publications into sub-collections of the archive, together forming the domain structure (the map). Apart from the specific features of different methodologies, it will depend on the choice of the element to create these clusters (cited references, words or authors), as to what *kind* of structure we will have. In most cases a cognitive structure is preferred over a social structure. For this kind a co-word or co-citation–based structure fits best.

Within a science policy context a landscape is often seen as an interesting representation, but the utility is not always clear. Apart from the user not always understanding the map, he does not know how to interpret the structure as such. Most maps are in a snapshot of a certain moment (year or year block). What the user needs is a reference point in order to be able to see the meaning of this snapshot within a context. A film of the changing landscape with certain reference points enhances these structures seriously. This is what can be termed science dynamics.

But there is another aspect to be taken into consideration. If we are to provide an instrument for monitoring and evaluating scientific research within the science policy context, we should give much attention to the actors. They constitute the only element in the model in Figure 10.1 that may be affected directly by policy. Every strategy or decision will apply to the actors, not to language, nor to the archive. This means that any mapping study in a science policy context should integrate information or linkages to actors, at any desired level of aggregation.

For the objective of a study the stage of display appears also to be very important. Within the science policy context the actual question to be answered should be explicit, so that a proper interface to explore the map and underlying data can be designed. This design may differ considerably from a design of an IR application. Noyons, Moed, and van Raan (2001), Boyack and Börner (2002), and Boyack, Wylie, and Davidson (2002) are the only cases in which the interface is explicitly designed as a science policy or research management tool. It should be mentioned, however, that many

[1] As discussed in detail in De Looze and Lemairie (1997) and Noyons (1999) the stage of field delineation is of course the most important one. We should be confident about what data are going to be used to create the domain map.

applications (c.f., Chen, 2003) do refer to potentials within a policy context, but are not elaborated upon as such. In sum, in policy–related mapping studies the following issues appear important:

– Comprehensive overview;
– Reference to actual/present situation;
– Points of reference to interpret results;
– Dynamics of the structure;
– Actors behind the structure;
– Actors behind dynamics.

As far as the map is concerned we need a limited amount of elements to be plotted in order to be able to provide a comprehensive overview. Most applications discussed in the previous section provide complex networks of many elements. A clustering of elements could in most cases reduce the amount of information in the map considerably. These clusters are referred to as paradigms (Small, 2003) or themes (Callon, 1983). This reduction of items in a map improves the stability and possibility of having points of reference and interpretation.

The reference to the actual situation (the real world) should be covered by the labelling of map elements whilst, of course, these labels should represent the contents of the (clustered) elements. It seems this issue has not had much attention yet, and depends heavily on expert input. For this the expert needs to recognize the structure and needs to be able to understand the individual elements identified.

Moreover, there are points of reference needed in order to interpret results. In the case in which we use a map to characterize the activity or impact of an actor, we need to know how this outcome relates to some point of reference, e.g., a world average, a similar actor or different period in time.

With respect to the dynamics it should be possible to visualize the changing landscape. This should, however, be established in such a way that changes can be monitored and interpreted easily. This means that for every change there should be a point of reference. This could be the starting or final point of a time series. In Noyons and Van Raan (1998) a method is outlined to deal with this.

As the elements in the map represent publication collections, and as actor data is in the author and affiliation fields of these publications, we are able to use the maps to distribute activity (numbers of publications per map element) and performance (activity and impact) over the landscape. Moreover, we should be able to integrate additional author information into the maps, such as funding and grants (c.f., Boyack and Börner, 2003). The interface should be able to enhance the map with this information, depending upon the question the user (science policy user) would like to have an answer to. Typical policy–related questions are:

– What does this domain look like?
– Who are the main actors in this domain?
– What is their particular expertise?
– How does this expertise relate to that of others?
– What are the main developments in a certain period of time?
– Which actors contribute to these developments?
– Who may be responsible for a particular change?

With these requirements in mind, many years were needed to develop at CWTS an interface that should be able to serve as a tool to provide possible answers to this kind of questions. In an example of a dedicated publication collection it will be shown how this interface works. And I will show how this approach provides possibilities to deal with the requirements as mentioned for a domain map as a tool for science policy or research management.

6. A CASE STUDY

In this section I will take the case of our own field as defined by a collection of relevant publications. The collection of publications is based on suggestions in Börner, Chen, and Boyack (2003). The collection is referred to as the ARIST collection or domain, after the serial in which this field delineation was used. It covers all core publications in scientometrics, bibliometrics, and domain visualization. The search strategy used is in Table 10.1. Data are retrieved from the ISI databases (*Web of Science*) in the years 1996–2002. I realize that a significant part of research is published outside the scope of the *Web of Science*, but for the main points I wish to make in this contribution the results are considered illustrative rather than complete.

Table 10.1. Search string for the ARIST data collection

citation analys* or
cocitation or co-citation or
(co occurrence and (word or term)) or
co term or co word or coword or coterm or
science map* or mapping science (map* of science) or
semantic analys* or semantic index* or semantic map or
bibliometric* or scientometric* or
data visualization or (visualization of data) or
information visualization or (visualization of information) or scientific visualization

An asterisk (*) indicates a right-hand truncation

With this delineation, the number of 1913 publications is covered. The distribution over years in the period 1996–2002 is in the Table 10.2.

Table 10.2. Numbers of publications in the ARIST collection

Year	*Number of publications*
1996	234
1997	223
1998	233
1999	270
2000	310
2001	307
2002	336

With the methodology described in Noyons (1999), I selected 237 keywords to be used to create the structure of the field (or domain). The cluster analysis yielded 12 keyword clusters as a good solution to represent the structure. These clusters are defined for the entire period of time and the evolution of the clusters (sub-domains) are presented in Table 10.3. The names of the sub-domains (labels) are set by the most frequent keyword. In other words, a sub-domain may be defined by a collection of keywords, but is represented in the table and map by only one.

Table 10.3. Numbers of publications per ARIST sub-domain, by year

Sub-domain	*1996*	*1997*	*1998*	*1999*	*2000*	*2001*	*2002*
1— impact	55	40	44	59	56	66	68
2— visualization tool	11	11	13	12	17	13	16
3— semantic analysis	17	11	19	23	20	24	17
4— information visualization	25	33	31	44	63	68	50
5— cluster analysis	7	3	5	4	6	9	7
6— information retrieval	5	5	8	7	15	8	13
7— journal	39	28	33	47	43	47	43
8— patent citation analysis	1	2	2	3	6	3	1
9— latent semantic analysis	4	2	7	6	7	8	4
10— bibliometric indicator	25	22	22	26	34	24	22
11— data visualization	12	14	12	14	18	16	33
12— citation analysis	26	21	20	30	24	28	26

It appears that in Table 10.3 most sub-domains remain at a similar level of activity during the entire period. In all sub-domains there is an increase of activity but it never exceeds the 20% growth on average during the entire period.

The overlap between the clusters (or sub-domains) is used as input to draw the map. In this map (Figure 10.2) we see 12 sub-domains with a wide range of sizes. The label of each sub-domain is represented by the most frequent keywords. The most prominent sub-domains are *semantic analysis* (which contains a substantial portion of general research as well) *citation analysis* (amongst which are co-citation studies), *journal* (representing all journal based studies), and *information visualization* and bibliometric indicator (containing performance evaluation studies). The other sub-domains represent more specific work. This map shows much overlap with the Kohonen map (ET-Map) presented in Börner, Chen, and Boyack (2003).

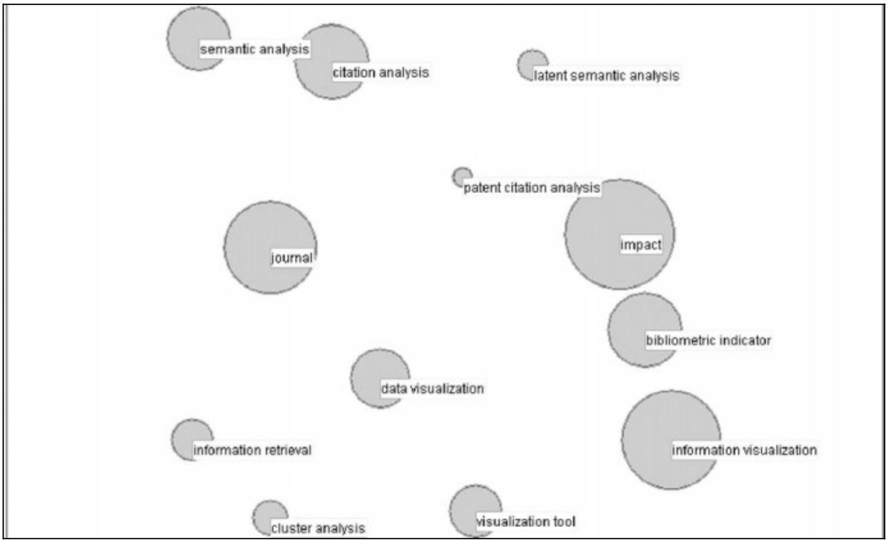


Figure 10.2. Map of the ARIST domain based on keywords co-occurrence

Separately we compiled a list of most active countries in the ARIST domain in the period 1996–2002. For these countries I also calculated the standard CWTS indicators, described in Moed, De Bruin, and Van Leeuwen (1995). The impact figures were at the time of the analysis available only for the period 1996–1998. The indicators show the profile per country for the entire field but in certain cases it may be interesting to see in which themes these countries are actually working on.

In other words, if we take the landscape as shown in the map, it would be interesting to know where the activity of these countries is located? I will illustrate how this information can be used to underpin these overall indicators. We take the example of the European Union (EU15, here considered as a country while all individual EU15 countries are included in the list as well) and compare it to the results for the United States (US). The US and EU15 have a similar production in this period but show some differences with respect to the other indicators. These differences may be due to the fact that they focus on different areas in the map. I will not elaborate on these differences here.

Table 10.4. List of most active organizations identified in the ARIST domain (1995–2002)

Country	*P*	*CX*	*CPP*	*PS*	*PN*	*CPP/FCSm*
USA	904	1,097	4.12	21%	8%	1.7
EU15	824	678	2.59	31%	12%	1.0
UNITED KINGDOM	193	187	3.53	32%	9%	1.3
GERMANY	185	92	1.70	30%	13%	0.6
NETHERLANDS	127	195	4.43	25%	5%	2.5
SPAIN	79	45	1.61	28%	17%	0.3
FRANCE	78	47	1.47	45%	20%	0.7
CANADA	76	46	1.64	36%	18%	0.7
JAPAN	63	38	1.73	38%	20%	0.6
ITALY	48	19	1.36	44%	14%	0.6

Country	Country or EU.
P	Number of publications.
Cx	Number of citations, self-citations excluded.
CPP	Average number of citations per publication (impact).
PS	Percentage of self-citations.
PN	Percentage of non-cited publications.
CPP/FCSm	Impact normalized by world average.

The map in Figure 10.3 below shows the distribution of the US over the ARIST domain. The shading of sub-domains is based on the relative contribution of the actor. Dark grey sub-domains are those with a high US contribution relative to their average contribution to the field. It appears that the US is present in all sub-domains, but with a clear focus on the upper (semantics) and lower part (visualization and IR) of the map.

The map in Figure 10.4 below shows the activity distribution of the EU15 in the ARIST Domain. Their cognitive orientation differs from that of the US, in the sense that the EU countries appear to focus more on the middle area of the map (journal studies and research evaluation). Apart from the fact that with these major actors (US and EU) we may expect that they are counterparts of each other with respect to activity in different sub-domains, we discern this clear difference of focus. From the literature in the field it is known for a fact that in Europe there is a major interest in the research evaluation component of bibliometrics.

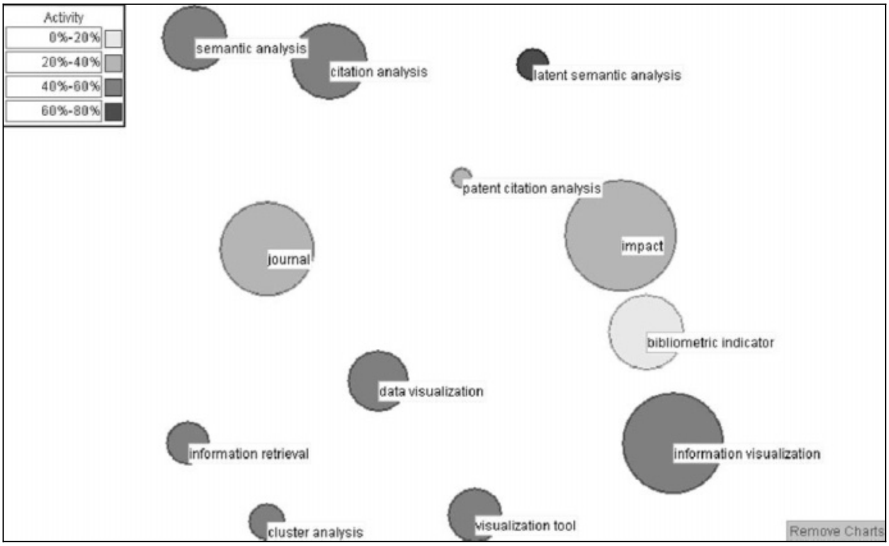


Figure 10.3. Activity profile of the US in the ARIST domain

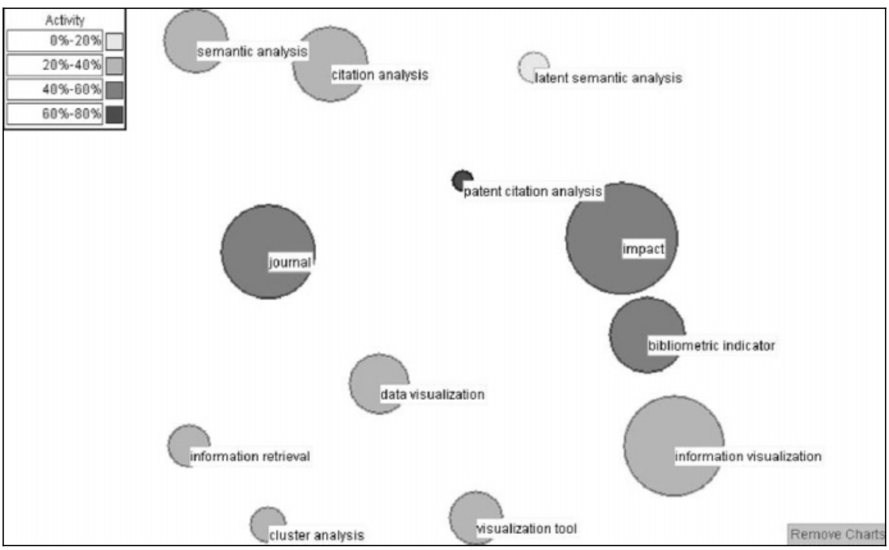
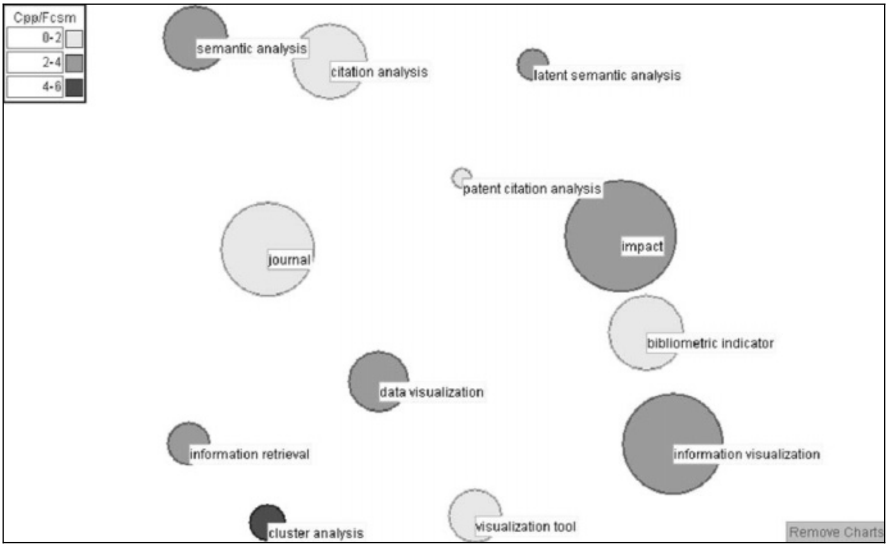
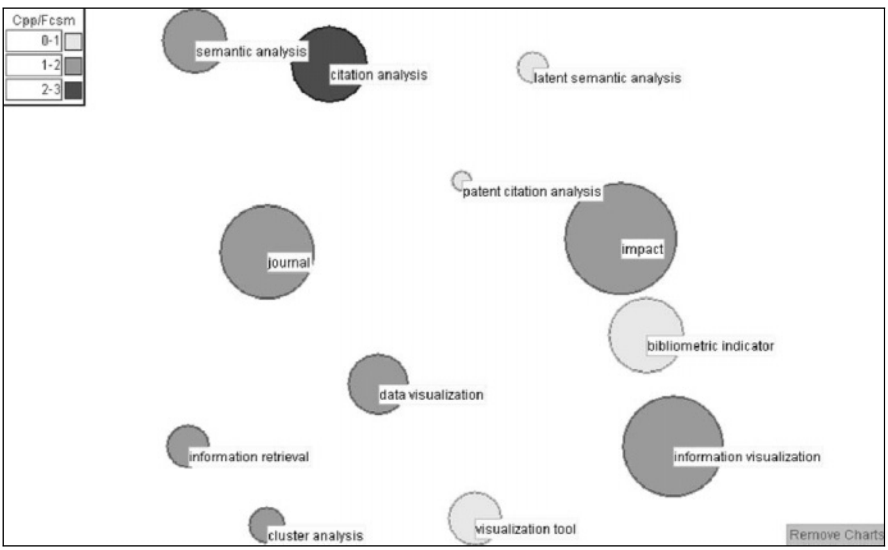


Figure 10.4. Activity profile of the EU-15 in the ARIST domain



United States



EU 15

Figure 10.5. Impact profiles of the US and EU-15 in the ARIST domain

The maps above (Figure 10.5) show that the impact distribution for both US and EU differs from their activity profile.

From a science policy perspective this distribution reveals the scope of activity of these two major actors. It would also be interesting to show their contribution to the field in terms of their impact. To investigate this I calculated the normalized impact (CPP/FCSm) of the publications for these two organizations and plotted them in the same map. This distribution shows where the impact of their work is located. Activity and orientation is one, the location of their impact is another. In this study I confine myself to showing the results. The number of publications in some cases is so low that it would be inappropriate to draw serious conclusions.

Because this small study is performed to illustrate the issues raised in this chapter, I left out detailed information about the analysis and results. More details are available via the author. This simple example illustrates how we can use the map and the actor information (organization names in publications) to characterize in more detail the activity and impact of actors and to relate it to an overall view of the domain (the map). This kind of displays provides a point of reference which makes it easier to interpret the results, because the same structure is used to characterize the performance of two different actors. In addition we could do these analyses using different time periods or additional indicators.

7. CONCLUSIONS AND PERSPECTIVES

In this chapter an overview is given of the history and present state of the mapping of science techniques, in particular those used within a science policy or research management context. The development of new techniques in the emerging field of *domain visualization* provides new opportunities. The idea of generating overviews of fields (or domains) is becoming more important in a time where fields become more interdisciplinary and where experts seem to have more problems maintaining this helicopter view.

As the domain visualizations (maps) and interfaces are dedicated mostly to IR objectives, the specific requirements for science policy and research management use should become explicit. Hence, we will be able to develop the proper maps and interfaces.

In view of the utility for science policy, the visualization techniques should provide points of reference to interpret the results more accurately. These reference points should apply to all elements at stake in such studies: structure; actors; and indicators.

The case study based on the technique developed and used at CWTS is used to illustrate the present situation and discuss the issues put forward in her. We hope to be able to join forces with many other domain visualization techniques to make serious progress in the analytical stages of the bibliometric mapping studies. We believe that the design of the user interface is of crucial importance to the applicability and utility for specific purposes, such as science policy and research management support.

REFERENCES

Ahlgren, P., Jarneving, B., Rousseau, R. (2003). Requirements for a co-citation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology,* 54, 550–560.

Bhattacharya, S., Basu, P.K. (1998). Mapping a research area at the micro level using co-word analysis. *Scientometrics,* 43, 359–372.

Börner, K., Chen, C., Boyack, K.W. (2003). Visualizing knowledge domains. In: Blaise Cronin (Eds.), *Annual Review of Information Science and Technology* (*ARIST*), 37, 179–255.

Boyack, K.W., Borner, K. (2003). Indicator–assisted evaluation and funding of research: visualizing the influence of grants on the number and citation counts of research papers. *Journal of the American Society for Information Science and Technology,* 54, 447–461.

Boyack, K.W., Wylie, B.N., Davidson, G.S. (2003). Domain visualization using VxInsight for Science and Technology Management. *Journal of the American Society for Information Science and Technology,* 53, 764–774.

Cahlik, T. (2000). Comparison of the maps of science. *Scientometrics,* 49, 373–387.

Callon, M., Courtial, J.P., Turner, W.A., Bauin S. (1983). From translations to problematic networks: an introduction to co-word analysis. *Social Science Information,* 22, 191–235.

Callon, M., Law, J., Rip, A. (1986). *Mapping the dynamics of science and technology*. London: The MacMillan Press Ltd. ISBN: 0 333 37223 9.

Chen, C. (2003). *Mapping scientific frontiers: the quest for knowledge visualization*. London: Springer Verlag. ISBN 1-85233-494-0.

De Looze, M.A., Lemarie, J. (1997). Corpus relevance through co-word analysis: An application to plant proteins. *Scientometrics,* 39, 267–280.

Garfield, E., Pudovkin, A.I., Istomin, V.S. (2003). Why do we need algorithmic historiography? *Journal of the American Society for Information Science and Technology,* 54, 400–412.

Kohonen, T. (1990). *The Self–Organizing Map*. In Proceedings of the IEEE, 78 (9), 1464–1480.

Kostoff, R.N., Eberhart, H.J., Toothman, D.R. (1998). Database tomography for technical intelligence: a roadmap of the near–earth space science and technology literature. *Information Processing and Management,* 34, 69–85.

Lin, X., White, H.D., Buzydlowski, J. (2003). Real–time author co-citation mapping for online searching. *Information Processing and Management,* 39, 689–706.

Merton, R.K. (1942). Science and technology in a democratic order. *Journal of Legal and Political Sociology,* 1, 115–126.

Moed, H.F., de Bruin, R.E., van Leeuwen, Th.N. (1995). New bibliometric tools for the assessment of national research performance: database description, overview of indicators and first applications. *Scientometrics,* 33, 381–442.

Noyons, E.C.M. (1999). *Bibliometric Mapping as a Science Policy and Research Management Tool*. Leiden: DSWO Press.

Noyons, E. (2001). Bibliometric mapping of science in a science policy context. *Scientometrics,* 50, 83–98.

Noyons, E.C.M., van Raan, A.F.J. (1998). Monitoring scientific developments from a dynamic perspective: self–organized structuring to map neural network research. *Journal of the American Society for Information Science,* 49, 68–81.

Noyons, E.C.M., Moed, H.F., van Raan, A.F.J. (1999). Integrating research performance analysis and science mapping. *Scientometrics,* 46, 591–604.

Persson, O. (1994). The intellectual base and research fronts of JASIS 1986–1990. *Journal of the American Society for Information Science,* 45, 31–38.

Persson, O. (2001). All author citations versus first author citation. *Scientometrics,* 50, 339–344.

Peters, H.P.F., van Raan, A.F.J. (1991). Structuring scientific activities by co-author analysis: an exercise on a university faculty level. *Scientometrics,* 20, 235–255.

Polanco, X., Francois, C., Lamirel, J.C. (2001). Using artificial neural networks for mapping of science and technology: A multi-self-organizing-maps approach. *Scientometrics,* 51, 267–292.

Salvador, M.R., Lopez-Martinez, R.E. (2000). Cognitive structure of research: scientometric mapping in sintered materials. *Research Evaluation,* 9, 189–200.

Schwechheimer, H., Winterhager M. (2001). Mapping interdisciplinary research fronts in neuroscience: A bibliometric view to retrograde amnesia. *Scientometrics,* 51, 311–318.

Small, H. (2003). Paradigms, Citations, and Maps of Science: A personal history. *Journal of the American Society for Information Science and Technology,* 54, 394–399.

White, H.D. (2003). Author co-citation analysis and Pearson's r. *Journal of the American Society for Information Science and Technology,* 54, 1250–1259.

White, H.D., Griffith, B.C. (1981). Author co-citation: a literature measure of intellectual structure. *Journal of the American Society for Information Science,* 32, 163–171.

White, H.D., McCain, K.W. (1997). *Visualization of literatures*. In M.E. Williams (Ed.), Annual Review of Information Science and Technology (ARIST), 32, 99–168.

White, H.D., McCain, K.W. (1998). Visualizing a discipline: an author co-citation analysis of Information Science, 1972–1995. *Journal of the American Society for Information Science,* 49, 327–355.

Widhalm, C., Topolnik, M., Kopcsa, A., Schiebel, E., Weber, M. (2001). Evaluating patterns of co-operation: application of a bibliometric visualisation tool to the Fourth Framework Programme and the Transport Research Programme. *Research Evaluation,* 10, 129–140.

Wouters, P. (1998). The Signs of Science. *Scientometrics* 41, 225–241.

Ziman, J.M. (1984). *An introduction to science studies: the philosophical and social aspects of science and technology*. Cambridge: Cambridge University Press. ISBN: 0-521-34680-0.

Chapter 11

ANALYSING SCIENTIFIC NETWORKS THROUGH CO-AUTHORSHIP

Wolfgang Glänzel[1,2] and András Schubert[2]

[1] *Steunpunt O&O Statistieken, Katholieke Universiteit Leuven, Leuven, Belgium*
E-mail: wolfgang.glanzel@econ.kuleuven.ac.be
[2] *Institute for Research Organization, Hungarian Academy of Sciences, Budapest, Hungary*

Abstract: Co-authorship is one of the most tangible and well documented forms of scientific collaboration. Almost every aspect of scientific collaboration networks can be reliably tracked by analysing co-authorship networks by bibliometric methods. In the present study, scientific collaboration is considered both at individual and national levels, with special focus given to multinational collaborations. Both literature data and original results witnessed a dramatic quantitative and structural change in the last decades of the 20th century. The changes, to great extent, can be attributed to the universal tendencies of globalisation and the political restructuring of Europe. The standards and, particularly, the visibility of scientific research, as a rule, benefit from the ever increasing level of collaboration, but the profits do not come automatically. This fact underlines the necessity of a regular quantitative monitoring of inputs and outcomes, i.e., bibliometric surveys.

1. INTRODUCTION

Scientific collaboration is a complex social phenomenon in research that has been systematically studied since the 1960s. Increasing collaboration was reported by Smith (1958), Clarke (1964, 1967), Price and Beaver (1966), Patel (1973) and Heffner (1979) in the context of growing funding, that is, directly or indirectly by economic factors. According to deSolla Price (1963) massive funding is one of the characteristics of 'big science'; team work is another. Team work requires large personnel, which, in turn, is strongly dependent of the financial support available for the research.

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 257-276.

Besides the economic factors, intra–scientific factors (see, for example, studies by deB. Beaver and Rosen, 1978, 1979; Luukkonen et al., 1992, 1993), especially changing communication patterns and increasing mobility of scientists, are also influencing collaboration. These factors motivate co-operation in 'less expensive' areas such as pure mathematics and theoretical research in social sciences, too. deB. Beaver (2001) has expanded the above–mentioned notion of funding–caused collaboration by giving a list of 18 purposes for which people collaborate. This list includes, beyond the access to funding and equipment, among others, also access to expertise, speeding up progress, enhancing productivity, and reducing isolation.

According to Patel (1973), involvement in (collaborative) research is manifested by authorship and by what he calls sub–authorship. Sub–authors are persons the contribution of whom is acknowledged by the authors of the publication as substantial assistance. In a recent study, Laudel (2001) has shown on the basis of a sample of interviewed scientists that a major part of collaboration is not acknowledged either through a proper acknowledgement or through co-authorship. A large share of persons involved in the preparation of a scientific paper does thus not appear either as co-author or as a sub–author of the publication. Consequently the question arises of how far co-authorship and sub–authorship are an adequate measure of collaboration. The relationship between *contributors*, *co-authors* (and *sub–authors*) and *co-writers* can thus be interpreted as a chain of subsets where co-authors form just a subset of contributors and those scientists who are actually writing the publications are, in turn, a subset of contributors acknowledged as *co-authors* and *sub–authors*.

Also Katz and Martin (1997) have found many cases of collaboration that are not 'consummated' in co-authored papers. They argue that co-authorship is no more than a partial indicator of collaboration. Intensifying collaboration, however, goes with growing co- and sub–authorship, as has been shown in several studies (for instance, Patel, 1973). We can thus conclude that there is at least a positive correlation between collaboration and co-authorship and sub–authorship at the level of individual actors.

The phenomenon described by Laudel and Katz and Martin rather applies to so-called *intramural* collaboration, that is, to collaboration within one department, research group or institute. Extramural collaboration, above all international collaboration, on the other hand, is usually well acknowledged. Moreover, does the relationship between collaboration and authorship change if not individual authors, but supra–individual co-operative research, such as that of teams, institutes, or even countries, is considered? The answer is yes, although, the relationship at the intra–institutional level at the same time does not necessarily change.

Kretschmer (e.g., 1994) has analysed aspects of social stratification in scientific collaboration at the micro level. The main findings are that extramural collaboration is characterised by similarity of the social status, whereas intramural collaboration shows significant differences of the social status of the co-authors. Interpreting these findings in the context of Laudel's results, we can conclude that the contribution of a quite large number of co-workers with lower status is 'only' acknowledged through sub–authorship or might remain even unacknowledged.

Newman (2001, 2003, 2004) has analysed the structure of collaboration networks on the basis of individuals' co-authorship patterns. Amongst other things, he has shown that co-authorship networks form 'small worlds' in which pairs of scientists are separated by only a short path of intermediate acquaintances. A similar model was developed and analyzed by Barabási et al. (2002).

Katz and Martin (1997) have found a conceptual problem with the single–authored but multi-institutional, or even international, papers. The paper by Glänzel (2001) on national characteristics of international collaboration itself might serve as an example for such a form of 'collaboration' which is often caused by multiple assignment and/or mobility of the author. In any case, it manifests an official relationship and involvement of the two or more institutions or countries. According to Katz and Martin at least 5–15 percent of collaborative papers at a national level seem to involve this form of 'collaboration' caused by multi-institutional authors. Despite this phenomenon, co-authorship seems to reflect research collaboration between institutions, regions, and countries in an adequate manner. Results of collaborative research at this level published in co-authored papers can thus be analysed with the help of bibliometric methods.

The findings by Laudel and Katz and Martin result in the conclusion that collaboration of individual scientists and that of institutions or of even higher levels of aggregation, have to be clearly distinguished. Nevertheless, the analysis of individual co-authorship gives insight into structural changes of collaboration at this level, too.

Institutional collaboration can, in turn, be studied in two important aspects: the first one concerns collaboration between different research institutions disregarding their organisational type. We consequently speak about inter–institutional collaboration — and in terms of published research results — about inter–institutional co-authorship whenever at least two different research institutions have contributed and thus the address of at least two different institutions appear in the byline of the paper. The effect of possible association between the institutions on collaborative patterns, as well as the role of strong federal structures in several countries with autonomous regions, makes, however, universal and comparative large-scale studies difficult. Most studies of inter–institutional collaboration are, therefore, restricted to national or regional analyses (e.g., Gómez et al., 1995, Hicks and Katz, 1997). A study of domestic inter–institutional collaboration in Canada, Australia, and the UK has uncovered an interesting phenomenon, namely, that research cooperation decreases exponentially with the distance separating the collaborative partners (Katz, 1994).

Co-operation between different sectors such as university, industry, and government is studied as a second important type of 'extra–mural' collaboration. Within the framework of the Triple Helix model introduced by Etzkowitz and Leydesdorff (for instance, Leydesdorff and Etzkowitz, 1996) this type of collaboration has gained a new dimension.

As indicated above, institutional collaboration is shaped by institutional sectors (cf., scientific co-operation between universities and between firms, respectively) and collaboration across sectors is characterised by regional or national peculiarities. Similarly to the inter–institutional collaboration, studies of collaboration across sectors are preferably studied in their national context (e.g., Hicks and Katz, 1997).

A third phenomenon has been studied by Cronin (2001). He characterises the extraordinarily large number of authors of single papers in several subfields of biomedical research and in high energy physics as 'hyper–authorship'. Indeed, publications with hundreds of co-authors affiliated with dozens of institutes in ten and more countries are no longer the exception to the rule. *Cronin* questions the possibility of fixing the degree of the individual co-authors' contribution to the paper since "to be an author is not necessarily to be a writer". Besides this phenomenon, an increasing number of multi-institutional or multi-national publications can be observed in other fields, too (e.g., de Lange and Glänzel, 1997, Glänzel and de Lange, 1997). Nevertheless, most of these publications cannot be considered in the context of hyper-authorship.

The present paper — as a consequence of the phenomena described above — will be focussed on the following three levels:

1. The *individual level*, that is, the level of individual scientists not aggregated to any unit like department, institution, region, or country;
2. The *cross-national level*, that is, collaboration between scientists with affiliation in different countries regarded as collaborating pairs; and
3. Multi-*national collaboration*, that is, collaboration between more than two countries taken into account explicitly.

The paper gives an overview of the development and application of indicators based on co-authorship, of their methodological background, and their use in research evaluation. In addition, the paper provides a review of co-authorship analysis focussing on relevant issues at the abovementioned levels of aggregation.

The bibliometric data used for this study are based on bibliographic data extracted from the 1980–2000 annual cumulations of the *Science Citation Index®* (SCI) of the *Institute for Scientific Information* (ISI — Thomson Scientific, Philadelphia, PA, USA). All papers recorded in the annual volumes as *article*, *letter*, *note,* or *review* were taken into consideration. At the level of authorship of individuals the number of all co-authors has been counted for *articles* and *notes* only.

The papers were assigned to countries based on the corporate address given in the byline of the publication. All countries indicated in the address field have thus been taken into account.

2. METHODS AND RESULTS

Research studies and reports on national and European science and technology indicators have recently presented figures reflecting intensifying scientific collaboration and increasing citation impact in practically all science areas and at all levels of aggregation (see, e.g., Narin and Whitlow, 1990; Narin et al., 1991; Moed et al., 1991; Glänzel, 1995; REIST–2, 1997; Glänzel et al., 1999). In the following three subsections this phenomenon will be studied at three important levels of aggregations.

2.1 The Individual Level

In a recent paper Persson et al. (2004) studied the effect of inflationary bibliometric values. In this context they have found, besides a growing number of publications, a strong increase of the number of active authors. In particular, the number of papers has grown by slightly more than one third (36%) between 1980 and 1998, and the number of authors have grown by almost two thirds (64%) in the same time span. Thus the authors concluded that the only possible interpretation of this tendency lies in a change in the patterns of documented scientific communication, and collaboration has changed in the last two decades, and that this tendency has inflationary features. The question arises of whether the density of co-authorship networks at this level has increased by forming stable teams of co-authors or if collaboration is, rather, characterised by the temporary creation of occasional links.

In a first step all research articles (document type: *article* or *note*) indexed in the 1980, 1990, and 2000 volumes of the SCI were analysed. The results are summarised in Table 11.1.

Table 11.1. Statistics on the co-authorship distribution in all fields combined in selected years

SCI Volume	*Share of single-authors papers*	*Co-author mean*	*Reciprocal of harmonic mean*
1980	24.8%	2.64	0.52
1990	15.7%	3.34	0.43
2000	10.7%	4.16	0.37

The share of single-authored papers in all fields combined continuously decreased. While in 1980 still about one quarter of all papers had only one single author, this share decreased to roughly 15% ten years later to reach the level of 10% in 2000. The average number of co-authors shows that this development reflects increasing multi-authorship. The average paper already has nowadays more that 4 co-authors. The change of the reciprocal of the harmonic mean reveals further interesting details: The average 'contribution' of a co-author reduced from about a half (0.52) in 1980 to slightly more than one third (0.37) in 2000. The increasing deviation of the harmonic means from the arithmetic mean indicates *increasing inequality* in the co-authorship distribution.

The breakdown by subject fields (cf., Glänzel, 2002) shows that all areas of science are characterised by intensifying collaboration associated with the increase of the share of multi-authored papers. In the medical fields the share of single-authored papers decreased from about 22% in the clinical and experimental specialties and about 15% in biosciences and biomedical research in 1980 to somewhat less than 9% and 6%, respectively, in the year 1998. The mean co-authorship grew from roughly 3 to 4.5 in clinical medicine and to 5 in biosciences and biomedical research. The corresponding share in chemistry halved from 15% to 7.5% — the average number of co-authors changed from 2.7 to 3.6 in this field. In mathematics, finally, a field that has always been a domain of individual scientists rather than that of teams, the share of single-authored papers dropped from roughly two thirds in 1980 to 40% in 1998. A mathematical publication in 1998 had two co-authors on an average compared with 1.4 co-authors in 1980. Similar patterns and developments have been observed in the social sciences, too. *Cronin* et al. (2003) reported an essential increase of collaboration and co-authorship in *psychology*, while philosophy is less affected by this tendency.

At the same time productivity of authors seem to increase. The question arises of how co-authorship and publication activity interact. In a recent paper by *Braun* et al. (2001) on publication and collaboration patterns in *neuroscience*, as well as in the above-mentioned study by Glänzel (2002), the interaction of co-operativeness and publication activity has been analysed. When average productivity is plotted against mean co-operativeness, field specific patterns can usually be observed: Productivity increases first with co-operativeness until a field specific threshold is reached; beyond this level, correlation turns negative. This threshold value ranges depending on subject peculiarities from 1–2 in mathematics, over 3–4 in chemistry, to 5–6 co-authors in neurosciences and biomedical research. These values beyond which collaboration does not exhibit higher productivity seem to be closely related to the co-authorship means of the corresponding fields (see the discussion above).

(Co-)authors can be classified into four types according to their anterior and posterior records. The relation between co-authorship and publication activity of the author types reveals information about the potential role of co-authors in forming stable teams or in creating occasional links. Price and Gürsey (1976) provided an elaborated scheme of what they called the "actuarial statistics of the scientific community". They introduced the categories *continuants*, *transients*, *newcomers*, and *terminators* which proved to be useful in the analysis of cooperation patterns (cf., Braun et al., 2001). According to the definition of Price and Gürsey, transients are authors publishing in the given year but neither before nor after, newcomers are authors publishing in and after the given year but never before, terminators were publishing before and in the given year but never after and continuants were publishing before, in and after the given year. For the continuants and, to a certain extent, also for the newcomers and terminators, a clear 'critical value' in the co-operativity–productivity plot as described above has been found in *neuroscience*. Productivity of transients is 'more uniformly' distributed over co-operativity without any distinguishable 'critical value'. The preference structure of authors of the four categories for cooperating with each other revealed another interesting aspect of co-authorship. The overwhelming part of papers co-authored by continuants and that co-authorship relations among these three categories, i.e., *transients*, *newcomers*, and *terminators*, are usually also mediated by continuants makes the notion of 'collaboration in stable teams' as the engine of intensification of co-authorship more than likely. This observation is in line with recent results by Newman (2004) on the structure of scientific collaboration networks.

The interaction of co-authorship with productivity is only *one* aspect of interaction with performance. A further important aspect can be analysed in the light of citation processes, namely, in the context of giving and receiving citations. Several recent papers have shown that collaboration has a measurable influence on citation behaviour. We just refer to the above-mentioned study by Persson et al. (2003) here. The authors have found that co-authored papers have longer reference lists than single-authored papers. Moreover, the number of references grow with the number of co-authors. Each co-author adds on average roughly half a reference to the list. The number of citations a paper receives is — on average — also strongly dependent of the number of co-authors. The effect is especially dramatic if authors from different countries have collaborated. However, this effect is subject to the following analysis, which will be deepened in the second section.

Summarising we can conclude that — even if co-authorship were indeed no more than a partial indicator of scientific collaboration at the level of individuals —, studying this phenomenon allows a deep insight into measurable interaction between collaboration and indicators of scientific communication and performance at this level, too.

2.2 Cross-National Collaboration

International co-authorship links have undergone dramatic structural changes in the last 25 years. Besides stable links and coherent clusters, new nodes and links in the international co-publication network have crystallised. In recent years fundamental mechanisms in international cooperation have been the subject of a number of different studies. In contrast to intra-national collaboration, where co-operation decreases with the distance of collaborating partners (see Katz, 1994), intensity of international collaboration is determined, besides by geographical proximity, by other factors, too. Among important factors influencing research collaboration are, e.g., the country size and political and economic reasons, as well as certain aspects of mobility and migration at the individual level. And unlike in the individual case where, indeed, those eighteen reasons listed by deB. Beaver (2001) are the main motivation why people collaborate in scientific research, there are also strong influences of historical, cultural and linguistic proximities on co-operation patterns at the national level. When one considers international collaboration the economic and/or political dependence of a country or geopolitical region (such as the different forms and degrees of neo-colonial ties) or large or special equipment (such as CERN in Switzerland and the observatories in Spain or Chile), which are often shared in large multinational projects, also condition co-operation, apart from any individual motivation. However, by far not all collaboration links between individual countries reflect symmetric relationship. Some links between several countries are thus characterised by specific unidirectional (or better asymmetric) co-authorship affinity (cf., Glänzel and Schubert, 2001). Some of these asymmetric patterns with strong historical background must be interpreted in the context of so-called *strong neo-colonial ties in science* (see Nagtegaal and de Bruin, 1994).

Some of the main lessons concerning international co-operation/co-authorship were concisely summarised in the comprehensive study of Narin and his co-workers (Narin and Whitlow, 1990; Narin et al., 1991). According to their key findings:

1. Internationally co-authored papers, whether co-authored by two countries within the EC or between an EC and a non-EC country, were twice as heavily cited as papers from a single EC country;
2. There was a steady rise in international co-authorship within and outside of the EC, and within and outside the EC targeted fields;
3. Tendencies for international cooperation were independent of country size and determined mainly by linguistic and historical factors;
4. The analysis of publications from the Less Favoured Regions of the EC revealed that the co-operating capabilities of these regions were very field dependent, corresponding in general to their national profiles.

International co-authorship, which is, in contrast to the level of individuals, assumed to reflect collaboration in a rather adequate manner, is accepted as a basically *positive phenomenon*. Nevertheless, Braun, Glänzel and Schubert have also pointed to problematic aspects of international collaboration. Thus extensive collaboration might be used as means for compensation for the negative financial effects which have hit the basic research system of several East European countries before and after the political and economical changes of the nineties (see Braun and Glänzel, 1996). The strong neo-colonial ties binding small scientific systems to those of large economies abroad might serve as another example. Strong (asymmetric) affinity thus may express a high degree of dependence of a scientific system from others. The continuously growing share of French co-publications with Algeria and Morocco in this context is striking.

Comprehensive macro-studies of international collaboration in the sciences by Schubert and Braun (1990) and Glänzel (2001) have shown that the share of internationally co-authored papers in most countries dramatically increased during the last two decades. In their study Schubert and Braun observed that foreign co-authorship can be approximated by national publication productivity through a power law in which the exponent is less than one. Big countries have thus, in general, lower shares of international co-publications than medium-sized or small countries have. Nevertheless, the growth of the share of international co-publications can be observed independently of the country's size. The increase is thus a global law. We will give some examples to illustrate this effect. According to Schubert and Braun the share of internationally co-authored papers in the USA, USSR and Japan in the begin of the 1980s of the last century lay significantly below 10%; by the end of the last century these shares reached and partly exceeded the value of 20%. For big and medium-sized developed countries, there was an increase from about 10%–20% in the early 80s to about 30%–50% at the end of the 90s (cf., Glänzel, 2001).

However, not only number and strength of several bilateral links has increased during the last decades; the whole network of international co-publications has undergone dramatic structural changes. In order to visualise this structural change we will map the co-authorship links of the most active countries in the world broken down by country pairs for the years 1980, 1990 and 2000. We have used Salton's (cosine) measure as an indicator of international collaboration strength. This measure is defined as the number of joint publications divided by the square root of the product of the number (i.e., the geometric mean) of total publication outputs of the corresponding pair of countries (cf. Glänzel, 2001). In order to guarantee that the results can be considered statistically reliable, we have chosen countries with at least 2000 publications in all fields combined in 1990 or 2000, and have plotted links for country pairs with at least 50 joined publications at three different levels of strength. The dramatic intensification of international co-publication as well as the structural changes in the collaboration network is presented in Figures 11.1 through 11.3.

The map presented in Figure 11.1 resembles those in Figure 11.3 in *Schubert* and *Braun* (1990) and in Figure 5 in Glänzel (2001). Schubert and Braun have analysed the international collaboration of 36 countries in the sciences in the period 1981–1985 and Glänzel has compared collaboration patterns of the most active 50 countries in 1985/1986 (Figure 5 in his paper) with those in 1995/1996. Little can be added to their findings. The authors detected several clusters of unequal size, namely, a big one including Western Europe, USA, and Canada, and two smaller ones with the Scandinavian and the Eastern European countries, respectively. Three tiny clusters, finally, included Australia and New Zealand, Egypt and Saudi Arabia, and Brazil and Argentina, respectively.

Germany and the USA can be considered the most important partners for the East European countries in the period around 1990. Interesting is the great share of Romanian–French co-authorship and the almost outstanding role of German co-operation in Bulgaria and Czechoslovakia. This confirms earlier results according to which Germany is usually the first important co-operation partner for East European scientific communities (see Glänzel and Winterhager, 1992); Germany can thus be considered the "gateway to the west" for Economies in Transition in Eastern Europe.

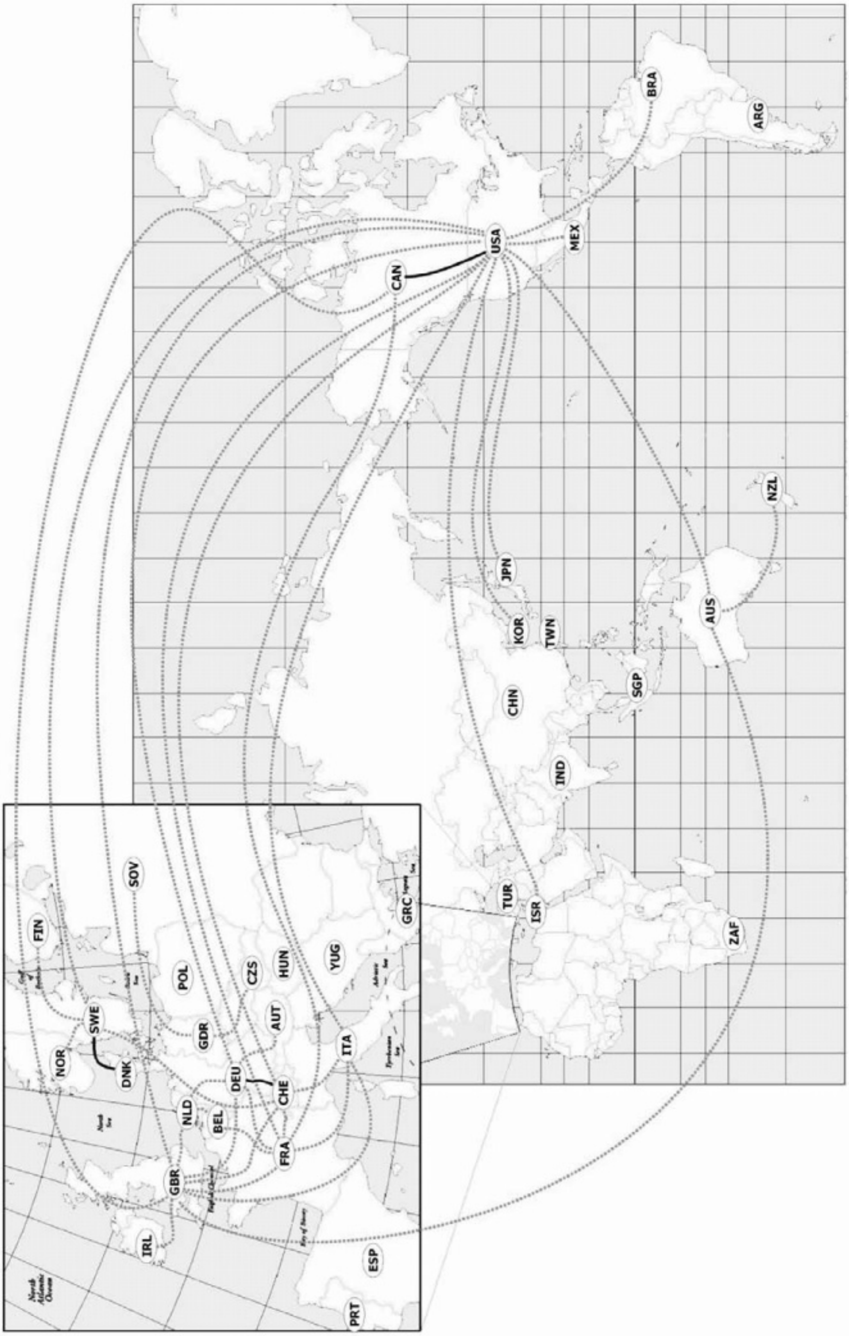


Figure 11.1. Co-authorship map for most active countries in all fields combined in 1980 based on Salton's measure (dotted line ≥ 1.0%, solid line ≥ 2.5%, thick line ≥ 5.0%)

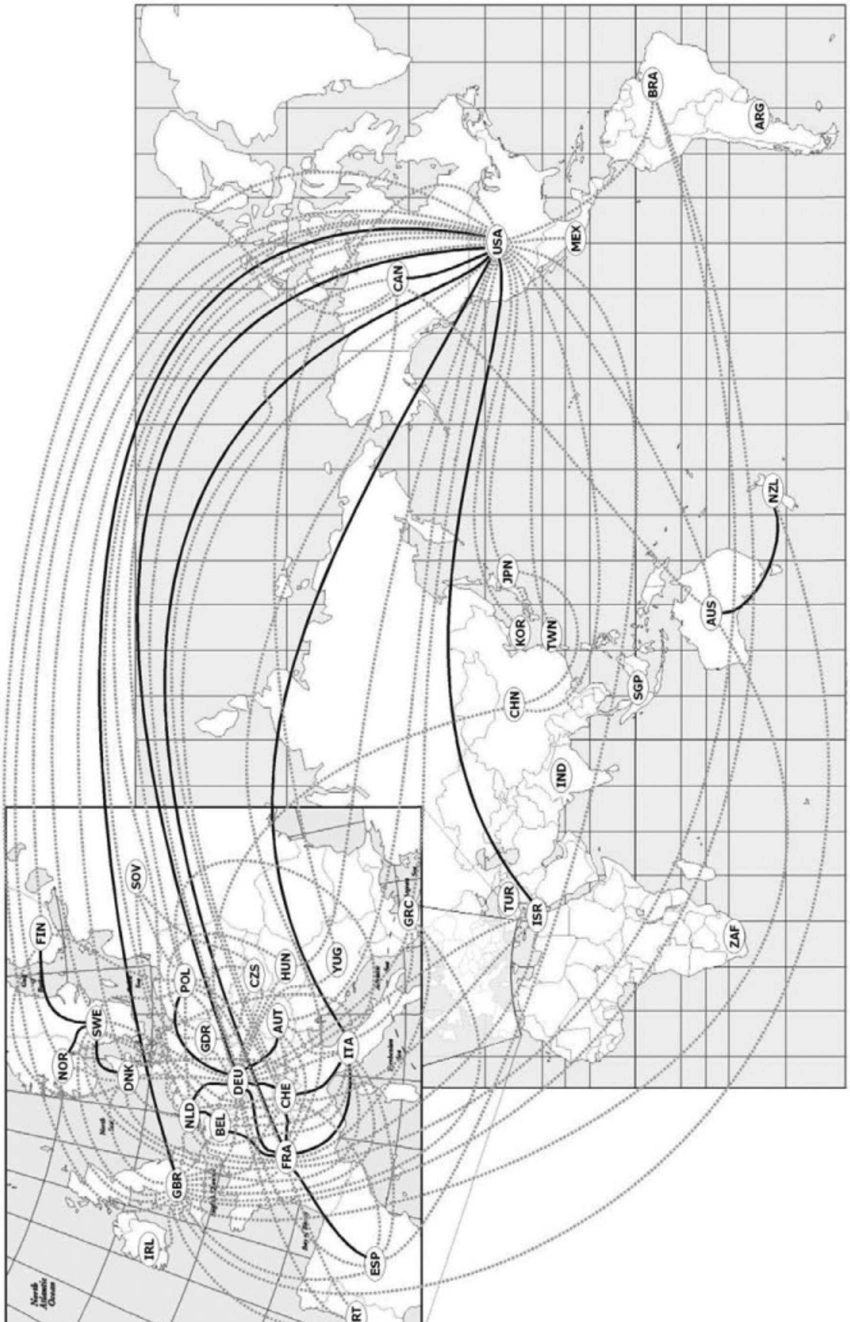


Figure 11.2. Co-authorship map for most active countries in all fields combined in 1990 based on Salton's measure (dotted line ≥ 1.0%, solid line ≥ 2.5%, thick line ≥ 5.0%)

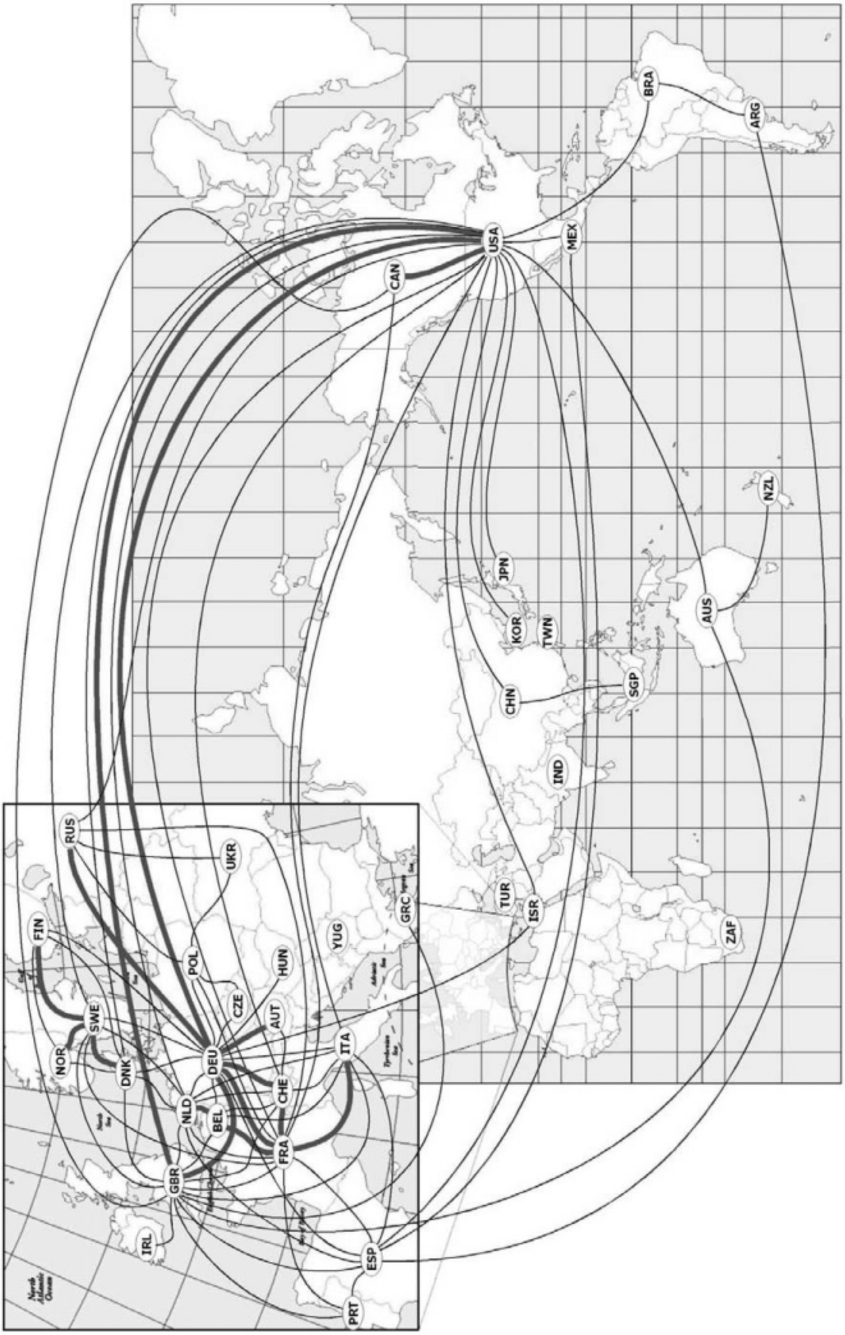


Figure 11.3. Co-authorship map for most active countries in all fields combined in 2000 based on Salton's measure (dotted line ≥ 1.0%, solid line ≥ 2.5%, thick line ≥ 5.0%)

From the global perspective Germany is besides the USA, France, and UK, also one of the world's most important nodes in the network of international collaboration (cf., Figures 11.2 and 11.3).

In addition to the above maps we would like to report interesting observations we have made concerning smaller countries not plotted in Figures 11.1–3. There are strong stable links both between Algeria and France, and Morocco and France which can be interpreted in the context of the above mention neo-colonial ties. Other local links arose and/or died out in the period of the last twenty years; amongst those we find the strong links between Cyprus and Bulgaria as well as Cyprus and Romania with strength of almost 8% according to Salton's measure in 2000. However, these links are based on small publication and co-publication set. A similarly strong link was established between Cuba and Mexico (6% in 2000). The strong link between the Czech Republic and Slovakia substantiate how closely the scientific systems of the two countries still are. The most interesting phenomenon is the co-publication link between China and Hong Kong: While there was practically no collaboration between these countries in 1980, their collaboration affinity evolved from 1.7% in 1990 to one of the strongest in the world map a couple of years before the crown colony returned to China (5.9% in 1995/1996) in order to vanish by 2000.

The development of country links mapped in Figures 11.1–3 clearly shows a trend towards a global network of scientific collaboration. The third section will be devoted to this phenomenon, namely, from the perspective of multi-nationality of collaborative research. Before we deal with this aspect of collaboration we still look at national profiles, visibility and citation impact of international co-publications.

If one compares national publication profiles of domestic and internationally co-authored publications, both theoretically, 4 possible types of changes in the profiles can be observed. Using the *Activity Index* suggested by Frame (1977), Glänzel (2001) has shown in his study about national characteristics in international collaboration that all of these four types occur in practice. In several countries collaboration has no significant effect on their publication profiles (e.g., Germany or Romania in 1995/1996), in other countries, such as Slovenia and Switzerland in the same period, collaboration seems to strengthen their national peculiarities. In a third group, for instance Japan and China, the opposite effect could be observed. In the last group the effect of collaboration on the national publication profile is less pronounced. Besides the well–balanced co-publication patterns of the first group, collaboration in the second and third group allows interesting speculations: Countries may conduct collaborative research mainly in their favourite subject field, or may, conversely, use collaboration as a means of compensating lacking domestic efforts in fields of relatively lower activity, which is, in turn, to the detriment of fields with higher activity.

The question arises of how far international collaboration has a measurable effect beyond national research profile also on *visibility* and *receptions*, i.e., on citation impact. The often observed relative high visibility and high citation attractivity of internationally co-authored publications resulted in what we can already consider the following commonplace: international co-publications appear in high impact journals and receive more citations than 'domestic' papers. On average this statement indeed seems to hold. The studies by Glänzel and Schubert (2001) and Glänzel (2001) have confirmed these assumptions. Nevertheless, the authors of these studies have shown that the above rule by far does not apply to all international papers.

In those fields where *targeting* is a more important aspect than 'global visibility', international collaboration often has a positive effect. In order to reach their audience, authors in clinical medicine often publish their results in their national language in their national journals. Their behaviour can, however, completely change if co-authors from abroad are involved. According to Glänzel (2001) the deviation of most countries' *mean expected citation rate* of international publications from that of domestic papers based on 3–year citation windows gauged against the world standard is in the fields clinical medicine, biomedical research, physics and engineering strictly positive. The opposite case, when a country is publishing its internationally co-authored papers on average in journals with lower impact can be observed, above all, in mathematics but, in part, also in chemistry and earth and space sciences. Not only are less advanced countries concerned, but also highly developed countries like the USA, in chemistry, and Australia, in mathematics, (cf., Glänzel, 2001). This observation may be at variance with the widespread notions concerning greater visibility of international co-publications. Similar contradictory observations could be made concerning the factually received citation rates of internationally co-authored papers. While the national totals of the citation impact of co-publications in all analysed fields often lie distinctly above the domestic 'standards', the situation changed if the citation impact is analysed by country pairs. Co-publications of several country pairs may attract fewer citations than expected on the basis of the corresponding domestic reference standards. Glänzel and Schubert (2001) called this type of co-publication links *cool links*. Unlike in biomedical research, where the observed citation impact of most analysed country pairs in the abovementioned study by Glänzel (2001) was higher than the domestic impact of at least one of the involved partners, and often higher than the world standard, too, the patterns in chemistry and mathematics reflect a somewhat different situation. Besides the outstandingly high citation impact of co-publications of several country pairs, the attractivity of joint papers of some pairs was unambiguously low in these fields. Here developing countries and Eastern Europe are the most concerned. International co-authorship seems, therefore, not always to pay for all partners. Glänzel has also analysed the citation distributions over domestic and international papers in his study, and found that citation patterns are normally characterised by significantly different frequency distributions and not merely, as one might expect, by statistical outliers such as a few highly cited or many uncited papers.

2.3 Multi-National Collaboration

The increase of multi-national co-authorship (besides the traditionally large multi-national research projects conditioned by special research facilities) reflects aspects of *globalisation* in scientific research.

Among the motivation for multi-national collaboration, the use of large equipment and instrumentation as a form of *big science*, economic reasons as well as political factors play an important part. Scientific collaboration is clearly stimulated (or hindered) by national, regional and global political interests. The auspices of the EC may serve as an example for a stimulation of the process of regional international scientific collaboration. Many multinational research projects are only possible in the framework of the large European programmes (Narin and Whitlow, 1990; Narin et al., 1991, REIST–2, 1997). Multiple institutional affiliations in the context with growing mobility have apparently become a further measurable factor. Nevertheless, multinational research — as all collaborative research — certainly not conditioned by these factors alone.

Figure 11.4 visualises the change of the share of bilateral, trilateral, and multinational international papers in time. 17% of all internationally co-authored papers in 2000 has authors with corporate addresses in more than two countries while this share amounted to 10% in 1990, and was still below 7% in 1980.

In order to measure the change of multi-national collaboration in time in the mirror of international co-publication links, de Lange and Glänzel (1997) and Glänzel and de Lange (1997) introduced the *Multilateral Collaboration Index*. This indicator is a national measure which can be interpreted as the mean number of partners involved in a country's international publications. According to de Lange and Glänzel (1997) and Glänzel and de Lange (1997), both the share of international papers and the degree of multilateral collaboration — mainly in the life sciences and in physics, but to a lesser extent also in chemistry and engineering — have grown considerably during the last decades. The values of the multilateral collaboration index have in practice increased for the most active 38 countries studied in the years 1983 and 1993.

In a recent study Glänzel and de Lange (2002) have analysed citation patterns of multi-national papers. They concluded that countries, in general, benefit from participation in multinational projects. This holds, again above all, for the life sciences. Nevertheless, in some cases in the natural sciences 'visibility' of multinational publications measured by 3–year journals impact did not significantly deviate from that of bilateral or even domestic papers.

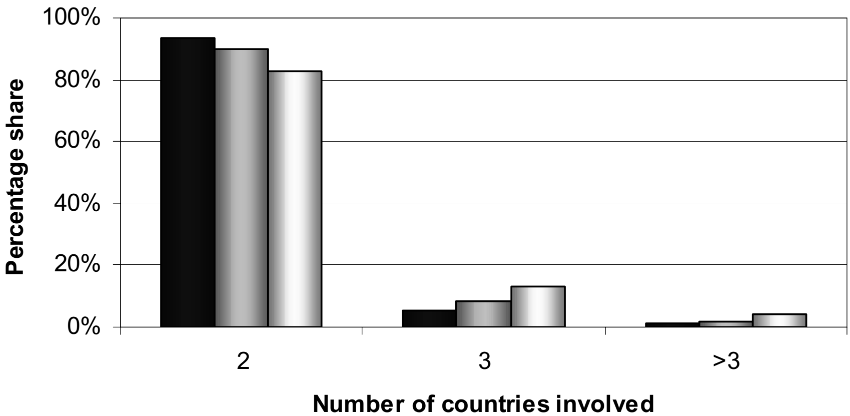


Figure 11.4. Share of countries involved in international publications (1980 — dark, 1990 — grey, 2000 — light)

It is perhaps worth mentioning that most of these multi-national co-authorships studied in the above papers by Glänzel and de Lange are based on collaboration of three or four countries, and thus cannot be considered as what *Cronin* (2001) has characterised as 'hyper-authorship'.

3. CONCLUDING REMARKS

Research collaboration and co-authorship in science is an interesting multi-faceted phenomenon. In order to understand and to interpret collaboration and co-authorship in a correct manner, co-operation must be studied at each level of aggregation in its specific way. Collaboration among individuals is at least in part subject to other motivations than collaboration between institutions and countries. Growing international collaboration is not only an expression of 'big science' but also part of the globalisation process in scientific research.

Nevertheless, whatever level we consider, research cooperation, in general, and co-authorship, in particular, appears to be 'cost effective', on the long run. It is true not only in economic sense, but also if more general value concepts, e.g., those measurable with bibliometric tools are concerned. Collaboration is able to promote research activity, productivity, and impact, and therefore is to be encouraged and supported by the means of research management and science policy. The benefits, however, do not come automatically. This fact underlines the necessity of a regular quantitative monitoring of inputs and outcomes, i.e., bibliometric surveys.

The results of the above research issues thus have beyond their significance for monitoring and mapping structural aspects of scientific research, also strong implications for the application of bibliometric indicators in research evaluation. Interaction of co-authorship with other important processes of scientific communication such as publication activity and citation behaviour may also result in reconsidering the construction of bibliometric standard tools to guarantee the validity of conclusions drawn from bibliometric results.

ACKNOWLEDGEMENT

The authors wish to thank Balázs Schlemmer for his creative assistance in preparing the figures of this paper.

REFERENCES

Barabási, A.L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A*, 311, 590–614.

Beaver, D. deB., Rosen, R. (1978). Studies in scientific collaboration. Part I. The professional origins of scientific co-authorship. *Scientometrics,* 1, 65–84.

Beaver, D. deB., Rosen, R. (1979). Studies in scientific collaboration. Part II. Scientific co-authorship, research productivity and visibility in the French elite. *Scientometrics,* 1, 133–149.

Beaver, D. deB. (2001). Reflections on scientific collaborations (and its study): Past, present and prospective. *Scientometrics,* 52, 365–377.

Braun, T., Glänzel, W. (1996). International collaboration: will it be keeping alive East European Research? *Scientometrics*, 36, 147–254.

Braun, T., Schubert, A., Glänzel, W. (2001). Publication and cooperation patterns of the authors of neuroscience journals. *Scientometrics,* 51, 499–510.

Clarke, B.L. (1964). Multiple authorship trends in scientific papers. *Science*, 143, 822–824.

Clarke, B.L. (1967). Communication patterns of biomedical scientists, *Federation Proceedings* 26, 1288–1292.

Cronin, B. (2001). Hyperauthorship: a postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*, 52, 558–569.

Cronin, B., Shaw, D., La Barre, K. (2003). A cast of thousands: Co-authorship and sub–authorship collaboration in the twentieth century as manifested in the scholarly literature of psychology and philosophy. *Journal of the American Society for Information Science and Technology,* 54, 855–871.

De Lange, C., Glänzel, W. (1997). Modelling and measuring multilateral co-authorship in international scientific collaboration. Part I. Development of a new model using a series expansion approach. *Scientometrics*, 40, 593–604.

Frame, J. D. (1977). Mainstream research in Latin America and the Caribbean. *Interciencia*, 2, 143–148.

Glänzel, W., Winterhager, M. (1992). International collaboration of Eastern Middle–European countries with Germany in the sciences, 1980–1989. *Scientometrics*, 25, 219–227.

Glänzel, W. (1995). International scientific collaboration in a changing europe. A bibliometric analysis of co-authorship links and profiles of 5 East–European countries in the sciences and social sciences, 1984–1993. *Science and Science of Science*, 4, 24–31.

Glänzel, W., De Lange, C. (1997). Modelling and measuring multilateral co-authorship in international scientific collaboration. Part II. A comparative study on the extent and change of international scientific collaboration links. *Scientometrics*, 40, 605–626.

Glänzel, W., Schubert, A., Czerwon, H.–J. (1999). A bibliometric analysis of international scientific co-operation of the European Union (1985–1995). *Scientometrics*, 45, 185–202.

Glänzel, W., Schubert, A. (2001). Double effort = double impact? A critical view at international co-authorship in chemistry. *Scientometrics*, 50, 199–214.

Glänzel, W. (2001). National characteristics in international scientific co-authorship. *Scientometrics*, 51, 69–115.

Glänzel, W. (2002). Co-authorship patterns and trends in the sciences (1980–1998). A bibliometric study with implications for database indexing and search strategies, *Library Trends*, 50, 461–473.

Glänzel, W., De Lange, C. (2002). A distributional approach to multinationality measures of international scientific collaboration. *Scientometrics*, 54, 75–89.

Gómez, I., Fernández, M.T., Méndez, A. (1995). *Collaboration patterns of Spanish scientific publications in different research areas and disciplines*. In M.E.D. Koenig and A. Bookstein (Eds.), Proceedings of the Biennial Conference of the International Society for Scientometrics and Informetrics (pp. 187-196). Learned Inf., Medford, NJ, 187–196.

Heffner, A.G. (1981). Funded research, multiple authorship, and subauthorship collaboration in four disciplines. *Scientometrics*, **3**, 5–12.

Hicks, D., Katz, J.S. (1997). *The changing shape of British industrial research*, STEEP Special Report N° 6, SPRU.

Katz, J.S. (1994). Geographical proximity and scientific collaboration. *Scientometrics*, 31, 31–43.

Katz, J.S., Martin, B.R. (1997). What is research collaboration? *Research Policy*, 26, 1–18.

Kretschmer, H. (1994). Coauthorship networks of invisible colleges and institutional communities. *Scientometrics*, 30, 363–369.

Laudel, G. (2002). What do we measure by co-authorships? *Research Evaluation*, 11, 3–15.

Leydesdorff, L., Etzkowitz, H. (1996). Emergence of a Triple Helix of University–Industry–Government Relations. *Science and Public Policy*, 23, 279–286.

Luukkonen, T., Persson, O., Silvertsen, G. (1992). Understanding patterns of international scientific collaboration. *Science, Technology & Human Values*, 17, 101–126.

Luukkonen, T., Tijssen, R.J.W., Persson, O., Silvertsen, G. (1993). The measurement of international scientific collaboration. *Scientometrics*, 28, 15–36.

Moed, H.F., De Bruin, R.E., Nederhof, A.J., Tijssen, R.J.W. (1991). International scientific co-operation and awareness within the European Community: problems and perspectives. *Scientometrics*, 21, 291–311.

Nagtegaal, L.W., De Bruin, R.E. (1994). The French connection and other neo–colonial patterns in the global network of science. *Research Evaluation*, 4, 119–127.

Narin, F., Stevens, K., Whitlow, E.S. (1991). Scientific co-operation in Europe and the citation of multinationally authored papers. *Scientometrics*, 21, 313–323.

Narin, F., Whitlow, E.S. (1990). *Measurement of scientific co-operation and coauthorship in CEC–related areas of science*, Volumes 1–2, Commission of the European Communities, Brussels – Luxembourg.

Newman, M.E.J. (2001). The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. USA*, 98, 404–409.

Newman, M.E.J. (2003). Coauthorship networks and patterns of scientific collaboration. *Proc. Natl. Acad. Sci. USA*, in press.

Newman, M.E.J. (2004). *Who is the best connected scientist? A study of scientific coauthorship networks*. To appear in E. Ben–Naim, H. Frauenfelder, Z. Toroczkai (Eds.), Complex Networks. Berlin: Springer.

Persson, O., Glänzel, W., Danell, R. (2004). Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies, *Scientometrics*, 60, forthcoming.

De Solla Price, D.J. (1966). *Little Science, Big Science*. Columbia Univ. Press, New York.

De Solla Price, D.J., Beaver, D. deB. (1966). Collaboration in an invisible college. *American Psychologist*, 21, 1011–1018.

De Solla Price, D.J., Gürsey, S. (1976). Studies in Scientometrics. Part 1. Transience and continuance in scientific authorship. *Internation Forum on Information and Documentation*. 1**,** 17–24.

REIST–2 (1997). *The European Report on Science and Technology Indicators 1997*. EUR 17639. European Commission. Brussels.

Schubert, A., Braun, T. (1990). World flash on basic research: international collaboration in the Sciences, 1981–1985, *Scientometrics*, 19, 3–10

Smith, M. (1958). The trend toward multiple authorship in Psychology, *American Psychologist*, 13**,** 596–599.

Chapter 12

PATENT CITATIONS AND THE ECONOMIC VALUE OF PATENTS

A Preliminary Assessment

Bhaven N. Sampat[1] and Arvids A. Ziedonis[2]

[1]*School of Public Policy, Georgia Institute of Technology, Atlanta, GA, USA*
E-mail: bhaven.sampat@pubpolicy.gatech.edu.
[2]*University of Michigan Business School, Ann Arbor, MI , USA*

Abstract Over the last decade, researchers studying innovation have increasingly used measures based on patent citations to estimate the values of new technologies, which are typically unobserved. In this study we examine the relationship between patent citation counts and private economic value in a dataset in which the latter is observed. Specifically, we use data about patenting and licensing by two major U.S. research universities to examine whether patent citations predict if university technologies are licensed, and the amount of revenue they earn if licensed. Our preliminary results suggest that citations are significantly related to the probability that a patent is licensed, but not to revenues conditional upon licensing.

1. INTRODUCTION

With the growing recognition of the economic importance of technological innovation, demand for measures of inventive outputs has increased dramatically over the past two decades. Measurement of such outputs has been frustrated, however, because key theoretical constructs such as ‘technological advance’ and ‘knowledge spillovers’ are not directly observable and thus difficult to quantify[1]. Because they are widely available

[1] Krugman (1991) has noted “knowledge flows … are invisible; they leave no paper trail by which they may be measured and tracked, and there is nothing to prevent the theorist from assuming anything about them that she likes” (p. 53). Measuring innovation is an important topic not only in the economics literature proper, but amongst scholars of the

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 277-298.

in electronic form, however, patent data have increasingly been employed to construct proxy variables for these unobserved concepts (e.g., Jaffe, 1998; Jaffe and Trajtenberg, 2002). The literature has made heavy use of patent citations (i.e., citations by patents to previous patents as prior art) as indicators of knowledge flows, and patent citation counts have been used as proxies for both the private and the social value of patented technologies.

In this study we investigate the relationship between revenues generated from the licensing of patents by two major research universities and the pattern of citations which these universities' patents receive. Our goal is to assess the degree to which patent citation counts are useful proxies for the private value of patented inventions. Validating citation based measures is difficult because of the dilemma posed by Trajtenberg et al. (1997), who ask "How can we establish the connection between a candidate proxy and [unobserved variable] x^*, given that by definition no direct data exist on x^*" (p. 31)[2]. Because so little is known about the relationship between patent citations and total economic value, we strive for a more modest goal — to examine whether citations can predict private value for a sample of patents where data on the latter (x^*) are available.

This chapter is organised as follows. Section 2 briefly discusses how scholars have interpreted patent citations and reviews previous validation studies of citations as economic indicators. In Section 3 we describe the dataset used in this study. Section 4 describes our econometric methodology and presents the main results. Section 5 concludes with a discussion of these results and implications for future research.

2. THE USES OF PATENT CITATIONS IN ECONOMICS

Patent citations and data based on patent citations have been employed in studies of innovation in response to limitations with the use of 'simple' aggregate patent counts as measures of innovative output (see Griliches, 1990 for a review). The large variance in the economic and technological significance of individual patents renders simple patent counts as extremely

management of innovation. Richard Rosenbloom recently suggested "It's the holy grail of people working on the management of technology—being able to measure innovation … I don't think anybody has cracked it, not yet" (quoted in Buderi, 2000).

[2] The authors point out that this question is not often asked in economics. See, however, Griliches (1983) and Klepper and Leamer (1984) on evaluation of proxy variables from an errors in variables framework, and Krasker and Pratt (1986) for a more general consideration.

noisy indicators of the innovative output of a firm or government program. A potential solution, suggested by Trajtenberg (1990), Narin et al. (1987), and Carpenter et al. (1981), was to weight patents by the number of times they are cited in subsequent patents.

Since that work economists have used citation counts to measure two different concepts of the 'value' of patented inventions. One is the patent's 'social value' based on the assumption that a patent B citing an earlier patent A draws upon on the knowledge embodied in A, or that A is a technological antecedent of B (Trajtenberg et al., 1997; Caballero and Jaffe, 1993; Jaffe and Trajtenberg, 1999). Citation of patent B by many subsequent patents suggests that numerous developments build upon the knowledge embodied in patent B (i.e., patent B has generated significant technological spillovers). Based on the rationale that inventions that generate a higher level of spillovers are more economically or technologically important, a significant stream of research has used counts of citations to patents to assess social value (or 'importance') of patented inventions (see Jaffe, 1998 for a review).

In many of the seminal articles, the 'building' metaphor is supported by analogy with bibliographic citations (citations in academic journal articles that indicate the sources upon which an author relies)[3]. Indeed, some suggest that patent citations are better indicators of sources upon which new knowledge builds than literature citations. Trajtenberg et al. (1997) argue that "because of the role of the examiner and the legal significance of patent citations, there is reason to believe that patent citations are less likely to be contaminated by extraneous motives in the decision of what to cite than other bibliographic data" (p. 23).

The 'building' metaphor may not be necessarily correct, however. Certainly, some of the previous patents upon which an invention builds (or technological antecedents, to use the language of Jaffe et al., 1998) will be "material to patentability" and cited, but not all will.[4] And there are certainly some items that a patent cites because they are "material to patentability" but upon which the invention did not build. These include not only citations inserted by examiners or the attorney that the inventor(s) did not know about when developing the invention (Sampat, 2003), but also very similar inventions which the inventor knew about but did not build upon or benefit from, cases where there is no economic spillover of the type discussed by Griliches (1992). Nevertheless, several recent papers by Adam Jaffe and

[3] This is an oversimplification, and there has been considerable debate in the sociology of science and bibliometrics as to the 'meaning' of bibliographic citations. See Melkers (1993) and Cole (2000) for recent overviews.

[4] Of course, many of the sources upon which inventions build are not patented.

colleagues (Jaffe, Trajtenberg, and Fogarty, 2000; Jaffe, Fogarty, and Banks, 1998) have found some support for the use of citation data for these purposes. However, more work remains to be done.

Scholars have also used citation counts as proxies for measures of the private value of the invention to the patent holder. Several studies have measured the value of a firm by its citation–weighted patent stock (Hall et al., 2000; Shane and Klock, 1997; Austin, 1994). Citations have also been used as measures of the value of inventions in studies of patent litigation (Lanjouw and Schankerman, 2000).

However, the literature is not always clear on how citations are related to the private value of innovations. In some analyses the implicit view is that social and private values of patented inventions are somehow correlated (Henderson et al., 1998)[5]. One way in which this could be true is if patent holders are able to appropriate a sufficient fraction of social returns to make citations a useful proxy for the private value of inventions. This reasoning suggests a first hypothesis regarding the relationship between patent citations and patent value:

Hypothesis 1: Citations represent the portion of social returns appropriated by the patent holder.

Trajtenberg (1990) proposes another mechanism through which citations may be related to the private value of patents, arguing that "the very existence of those later patents attests to the fact that the earlier patents opened the way to a technologically successful line of innovation. More important, they presumably attest to the economic success as well (at least in expected value terms), for those patents are the results of subsequent costly innovational efforts undertaken mostly by profit seeking agents ... If citations keep coming, it must be that the innovation originating in the citing patent had indeed proved to be valuable" (p. 189). In this view a patent that has been revealed to be profitable will induce other firms to undertake research in technologically close but non-infringing areas, (probabilistically)

[5] While the Henderson et al. (1998) discussion of the diminished 'quality' of university patents after, 1980 appears to reflect a social value construct (i.e., spillovers or the degree to which other inventions build upon university patents), the authors also suggest that a possible reason for this finding is that it was less costly for universities to patent after, 1980, which effectively lowered the threshold quality level of invention above which they would file for patent protection. If universities are comparing private costs with private benefits, as they appear to be in this explanation, then increased patenting would only occur if the 'quality' of the invention was correlated with the expected private value of the patent.

resulting in citing patents. This suggests a second hypothesis regarding the mechanism by which patent citations may represent private economic value:

Hypothesis 2: Citations reflect entry into profitable areas of research.

Similarly, citations may reflect the prior or current commercial interest and activity in a field, and hence profit potential, suggesting a third hypothesis:

Hypothesis 3: Citations indicate technological opportunities or market interest in a technological area.

In a study of patent litigation Lanjouw and Schankerman (2000) find that citations increase immediately after litigation, attributing the increased citations to a 'publicity effect'. They reason that once patents are litigated, they are better known to examiners and applicants, and thus more likely to be cited. This view suggests that patents that are economically successful are more widely known and more likely to be cited as prior art. A fourth hypothesis reflects this possible mechanism between citations and private value:

Hypothesis 4: Citations result from public disclosure.

In principle, detailed data about the timing of citations and identity of citers could help to distinguish which of the mechanisms represented by our four hypotheses (if any) are responsible for the citations–private value relationship. Hypothesis 1 (citations reflect social return which is correlated to private return) has implications for the identity of citers: it suggests that a significant share of spillovers are appropriated, and that these appropriated spillovers account for private value to the patent holder. In this view a significant share of citations may be generated by parties who are likely to compensate the patent holder (licensees), and such citations should be more closely related to economic value than citations by others. Hypotheses 2 (entry into areas of research revealed to be profitable) and 4 (publicity effect) link the timing of citations relative to realizations of private economic value, (licensing or profitability). In particular, these two hypotheses suggest that realizations of private returns will *precede* citations. Thus under Hypotheses 2 and 4 very early citations would not be correlated with private economic value, whilst later citation counts would be.

Using data about the timing and identity of the owners of patents that cite our sample of patents relative to the timing and identity of licensees of these same patents, we present some preliminary findings on the question of which of the four mechanisms presented above may explain the citations–value relationship. The main objective of this study, however, is to assess whether there is any relationship between citations and private value of individual patents at all. This exercise is similar in spirit to several recent attempts to validate citation counts as measures of private economic value. Some authors have used measures of value at the firm level, considering whether a firm's citation weighted patent stock appears to impact its market value (Deng et al., 1999, Shane and Klock, 1997; Austin, 1993; Hall et al., 2000). The most comprehensive of these studies is Hall et al. (2000), which finds a significant relationship between the 'Tobin's Q' of firms and their citation weighted knowledge stocks and other inputs. Other scholars have focused on the individual patent as the unit of observation. Lanjouw and Schankerman (1999), using a latent variable model, and find that citations are positively correlated with other measures of the value of patents.

In work more closely related to this study Harhoff et al. (1999a, 1999b) find that citation counts appear to reflect the 'asset value' of patents, or the price at which surveyed patent owners reported they would be willing to sell the rights to particular patents. In the only other study of which we are aware in which the authors had access to direct measures of private returns from patents, Mogee et al. (1997) compare citation counts to various measures of the value of a sample of patents, including an estimate of patent value from a patent renewal model, whether a patent is licensed, and the amount of license revenues earned by a patent. They find that the number of citations are positively and significantly related to renewal model value estimates and whether a patent is licensed, but that there is no significant relationship between citations and the level of revenues.

3. DATA

We utilise data about university patents, citations, license outcomes, and license revenues to examine the citations–value relationship. After the passage of the Bayh–Dole Act of, 1980, universities became more active in patenting and licensing inventions generated by faculty research (Henderson et al., 1998; Mowery et al., 2001). One advantage of using university data for studying the relationship between patent citations and economic value is that, unlike the private sector, the university lacks the requisite complementary assets and the motive to engage in product development and marketing activities to capture economic value. Universities typically also do not have strategic motives to patent to prevent competitors from using the new technologies or to block competitors from patenting. Therefore, universities typically apply for patent protection solely for the purposes of licensing inventions generated by research.[6] Another advantage to the university data is that few private firms are able to patent in as wide a range of technological fields as are major research universities. University patent data therefore also enable us to examine the relationship between citations and revenues across several broad technological fields. On the other hand, university patents may themselves be qualitatively different from corporate patents (Jaffe et al., 1993; Trajtenberg et al., 1997), which may limit the degree to which our results are generalisable. We revisit these issues below.

Table 12.1. Total Patents Assigned to the University of California and Columbia University, by Application Year

Application Year	*University of California*	*Columbia University*	*Total*
1980	55	2	57
1981	45	1	46
1982	50	5	55
1983	54	4	58
1984	57	9	66
1985	62	12	74
1986	60	12	72
1987	70	16	86
1988	76	21	97
1989	82	15	97
1990	72	12	84
1991	72	15	87
1992	76	17	93
1993	108	13	121
1994	91	24	115
Total	1,030	178	1,208

Our sample of patents was generated from archival data at the technology transfer offices of Columbia University and the University of California. These two universities are amongst the leading recipients of licensing revenues amongst U.S. research universities in recent years (AUTM, 2000). The Columbia and University of California data contain disclosures of inventions made by faculty, researchers, students, and staff at these two

[6] Mowery and Sampat (2001b) note, however, that earlier in the 20th century universities had other motivations for patenting as well, including preventing firms from patenting the fruits of university research and monopolising an emerging technological field; and to assure that only 'reputable' producers exploited university inventions, thus protecting universities from 'bad press'. As a result, university patents were often dedicated to the public. However, since the, 1970s, and especially since the passage of the Bayh–Dole act of, 1980, a primary motive for university patenting has been to license inventions.

universities, and the patenting and licensing outcomes of these invention disclosures. Our sample contains 1208 issued patents applied for by the University of California and Columbia University between 1980 and 1994. For both universities we observe whether each of these patents was licensed by the end of, 1999; and, if licensed, the identity of the licensee. For the Columbia sub-sample we also observe the total amount of dollar payments made to the university by that date. We also utilise the Micropatent database of US patents to identify all subsequent patents that cite our sample patents as prior art by the end of, 1999. Table 12.1 shows trends in total patents issued to the two universities by application year. The University of California (UC) accounts for 85% of the patents in our sample, reflecting its greater number of campuses (nine vs. two) and its long history of involvement in patenting — Columbia gradually entered into patenting activities only after the passage of the Bayh–Dole Act, in, 1980 (Mowery et al., 2001; Mowery, Sampat, and Ziedonis, 2001).

The dependent variables in our analysis are licensing outcomes: (a) whether a patent is licensed; and (b) the revenues it earns conditional on licensing. Table 12.2 shows the distribution over time of the first of these variables (whether licensed) for the pooled sample of patents. This table reports the proportion of patents in each application year that were licensed by, 1999. Within application years, and in the overall sample, fewer than half of the patents are ever licensed. Since licensing is the one of the only means through which universities reap returns from patents, this table suggests that less than half of university patents have the potential to earn revenues. This implies, furthermore, that the value of university patents will have a skewed distribution.

One empirical problem we face is that the series is right censored (i.e., we only know if a patent has been licensed by a given point in time, not whether it will be ultimately licensed). For example, although Table 12.2 reports that 41% of the patents in our sample applied for in, 1994 were licensed by, 1999, a greater fraction may eventually be. We consider this characteristic of the data in developing the econometric setup, and in the interpretation of our results.

Several issues complicate measurement of patent licensing revenue, however. First, patents are often licensed in bundles (‘inventions’), and revenues accrue to the inventions or groups of inventions rather than to individual patents. Because we are not able to observe the relative importance of each patent in a bundle, we have no rule for allocating licensing revenues across bundled patents. We therefore use licensed inventions as the unit of analysis in our licensing income specification.

Table 12.2. Distribution of Licensing Outcomes by Application Year

Application Year	*Proportion of Patents Unlicensed (%)*	*Proportion of Patents Licensed (%)*
1980	73.68	26.32
1981	76.09	23.91
1982	60.00	40.00
1983	58.62	41.38
1984	62.12	37.88
1985	54.05	45.95
1986	52.78	47.22
1987	50.00	50.00
1988	51.55	48.45
1989	58.76	41.24
1990	58.33	41.67
1991	56.32	43.68
1992	60.22	39.78
1993	61.16	38.84
1994	59.13	40.87
Total	58.69	41.31

A second issue concerns the types of revenues to include as 'license revenues'. Licensees may pay advance fees upon execution of the contract, annual fees to keep the license active, milestone payments based on level of sales or other events (e.g., reaching some stage of clinical trials), sales based royalties, as well as legal reimbursements and other fees. Since we are interested in the private value of the patented invention to the university, which is similar to the asset value of the patent to a firm lacking strategic motives (Harhoff et al., 1999), we use all revenues except for reimbursements.[7] A third issue is whether to treat unlicensed patents as observations with zero revenues or to exclude them from the analysis relating licensing income to citations. We chose the latter strategy both because the 'technology' unit of analysis is not properly defined for unlicensed inventions (patents are bundled at licensing) and because it allows for distinguishing between licensed inventions with zero revenues and unlicensed inventions.

Like the data on whether a patent is licensed, revenue data for the Columbia sub-sample are right–censored. They are therefore subject to truncation bias; at any given time we observe only a fraction of the lifetime revenues earned by licensed inventions. The resulting data on gross revenues for licensed technologies are also extremely skewed: in 1980 the top 10% of

[7] Additional analyses (unreported) show that the main qualitative results are not affected if we consider sales-based royalties alone.

licensed technologies account for 95% of total gross income, and in 1990 the top 10% account for 88%. This feature of the distribution, that outlying tail values account for a large proportion of cumulative revenue, is consistent with previous evidence on the distribution of returns from industrial innovations (see Scherer and Harhoff, 2000 for an excellent review), and university inventions (Mowery et al., 2001; Mowery and Sampat, 2001). As is common practice in studies employing highly skewed data therefore, we use a log-transformation of gross revenues as the dependent variable. Because many of the licensed technologies within the sample earn no revenues, we construct the dependent variable as the log of $1 plus license revenues. The patent citation distribution is also extremely skewed, as would be expected if patent citations were related to the value of inventions.

Table 12.3. Licensee Citations, Pre-License Citations for Licensed Patents by Application Year of Cited Patent

Application Year	*Proportion of Citations By Licensees (%)*	*Proportion Citations Occurring Before Licensing (%)*
1980	47	11
1982	13	15
1983	29	5
1984	9	2
1985	11	10
1986	7	18
1987	29	16
1988	3	40
1989	0	32
1990	14	19
1991	16	8
1992	15	12
1993	30	0
1994	15	0
1995	0	0
Overall	16	14

As mentioned earlier, one goal of this chapter is to shed light on the mechanisms by which citations and value may be related. Table 12.3 reports the citations made by licensees as a fraction of overall citations by application year of our sample patents. Overall, licensees account for a small share of citations (16%), although there is considerable variation over time.[8]

[8] Firms (especially multi-division firms) often hold patents in the names of their subsidiaries or their parent firms (see Hall et al., 2000). To ensure that we captured all citations made by licensees or their corporate parents or subsidiaries we utilised the *Directory of*

Since only licensee citations are associated with compensation to the university, this low number suggests that Hypothesis 1 — that citations reflect social value but that part of this value is appropriated by patent holders — is unlikely to be the main or only causal link between citations and value, at least in this sample. We reconsider this hypothesis by examining the relationship between licensee citations and revenues from licensed inventions below.

The third column of Table 12.3 shows that in general a low proportion of citations are made before the license execution date. This provides some support for hypotheses 3 and 4 (suggesting that citations follow economic success). We revisit this suggestion below.

Before turning to the econometric analysis it is useful to examine the basic relationships between the indicators of economic value and citations. Table 12.4 reports the mean number of citations for licensed and unlicensed patents. Both means decline over time, probably an artifact of truncation bias. In each application year however, licensed patents have a higher number of citations, on average, although the magnitude of the difference varies over time. Indeed, since both patent citation and the licensing series are truncated we can ask whether at a given point in time citations are informative of the license status of an invention or the revenues earned by an invention. Figure 12.1 shows a scatter plot of log revenues versus citations for technologies licensed by Columbia University. These data provide preliminary evidence that there is a positive relationship between revenues and citations, though the clusters of points along the x and y axes (uncited patents with revenues, cited patents without revenues) also suggest that the relationship is noisy.

Corporate Affiliations to identify the corporate affiliations of the assignee of each citing patent. If the licensee was a subsidiary or parent of the citing firm we considered that citation as being made by the licensee.

Table 12.4. Mean No. of Citations for Unlicensed, Licensed Patents

Application Year	*Mean No. of Citations for Unlicensed Patents*	*Mean No. of Citations for Licensed Patents*
1980	9.67	26.07
1981	7.17	15.55
1982	8.91	15.18
1983	12.44	16.42
1984	7.68	20.16
1985	10.40	13.74
1986	8.58	13.88
1987	7.91	8.84
1988	8.38	12.09
1989	5.72	7.00
1990	4.37	9.26
1991	3.96	6.82
1992	3.57	3.84
1993	1.66	2.11
1994	1.81	2.32
Overall	6.16	9.81

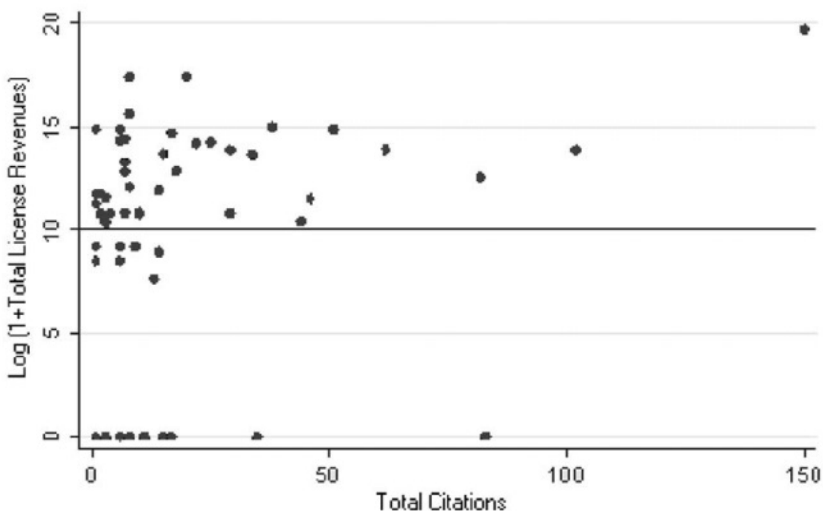


Figure 12.1: Scatterplot of Log Revenues vs. Citations, Columbia Technologies

4. ECONOMETRIC METHODOLOGY AND RESULTS

To examine whether citations can predict whether a technology is licensed, we estimated probit models regressing a dummy variable indicating if the patent was licensed on citations, technological field controls, and application year controls. Results are reported in Table 12.5. The positive and statistically significant coefficient for total citations (Model 1) indicates that citations are related to whether a technology is licensed. Probit coefficients give the effect of a one unit change in X on the cumulative normal probability of Y, and are thus difficult to interpret directly. We therefore calculate the marginal effect of citations on the probability of licensing (calculated at the mean of the data). The marginal effect is approximately 0.007 (i.e., an additional citation leads to a 0.7% increase in the probability that a patent is licensed). Figure 12.2 plots the predicted marginal effects across the range of the citations variable: note that the marginal effects are greatest (about 0.74%) at about 20 citations.

Because it is not strictly correct to speak of infinitesimal changes in integers such as citation counts (Caudill and Jackson, 1989) we also calculate the predicted probabilities of a patent being licensed as a function of the number of citations. At the mean level of citations (8 citations), the predicted probability of licensing is 41.2%, and at the median (4 citations) the predicted probability is 35.4%. Increasing citations from the median level to the 75th percentile level (9 citations) increases the probability of licensing by 3.6%, and increasing citations from the median to the 95th percentile level (28 citations) increases the probability by 17.6%. Figure 12.3 shows the predicted probabilities of licensing by the number of citations a technology receives.

Table 12.5. Probit Estimation for University of California and Columbia University Patents

	Dependent Variable: "Was the Patent Licensed?" (1=Yes, 0=No)			
	(1) Entire Sample	(2) Univ of California Patents Only	(3) Columbia Patents Only	(4) Entire Sample, Incl. Interaction Term
Total Citations	0.019 (5.36)**	0.017 (4.20)**	0.030 (3.54)**	0.019 (4.63)**
Citations x Columbia Dummy				0.000 (0.05)
1981 Dummy	0.014 (0.05)	0.021 (0.08)		0.014 (0.05)
1982 Dummy	0.356 (1.39)	0.387 (1.47)	0.951 (0.47)	0.357 (1.39)
1983 Dummy	0.418 (1.65)	0.298 (1.15)		0.418 (1.65)
1984 Dummy	0.303 (1.23)	0.202 (0.79)	2.323 (1.13)	0.304 (1.23)
1985 Dummy	0.525 (2.20)*	0.548 (2.21)*	1.688 (0.83)	0.525 (2.20)*
1986 Dummy	0.575 (2.39)*	0.682 (2.71)**	1.450 (0.71)	0.576 (2.39)*
1987 Dummy	0.705 (3.03)**	0.746 (3.07)**	1.849 (0.90)	0.706 (3.03)**
1988 Dummy	0.670 (2.93)**	0.573 (2.39)*	2.530 (1.24)	0.670 (2.92)**
1989 Dummy	0.531 (2.32)*	0.625 (2.64)**	1.212 (0.59)	0.531 (2.31)*
1990 Dummy	0.516 (2.20)*	0.522 (2.14)*	1.965 (0.96)	0.517 (2.19)*
1991 Dummy	0.593 (2.54)*	0.558 (2.29)*	2.238 (1.09)	0.594 (2.54)*
1992 Dummy	0.531 (2.28)*	0.492 (2.02)*	2.204 (1.07)	0.532 (2.28)*
1993 Dummy	0.520 (2.32)*	0.528 (2.29)*	1.918 (0.93)	0.521 (2.31)*
1994 Dummy	0.553 (2.45)*	0.614 (2.60)**	1.865 (0.91)	0.554 (2.44)*
Chemicals Dummy	-0.162 (0.68)	-0.188 (0.76)	-0.034 (0.04)	-0.162 (0.68)
Drugs and Medical Dummy	0.075 (0.33)	0.064 (0.27)	0.013 (0.01)	0.075 (0.33)
Electronics Dummy	-0.080 (0.34)	0.009 (0.04)	-0.809 (0.84)	-0.080 (0.33)
Mechanical Dummy	-0.676 (2.51)*	-0.637 (2.26)*	-1.076 (1.02)	-0.676 (2.51)*
Constant	-0.796 (2.70)**	-0.774 (2.55)*	-2.295 (1.02)	-0.798 (2.70)**
Nr Observations	1205	1027	173	1205

Absolute value of *z*-statistics in parentheses. * Significant at 5% ; ** Significant at 1% level.

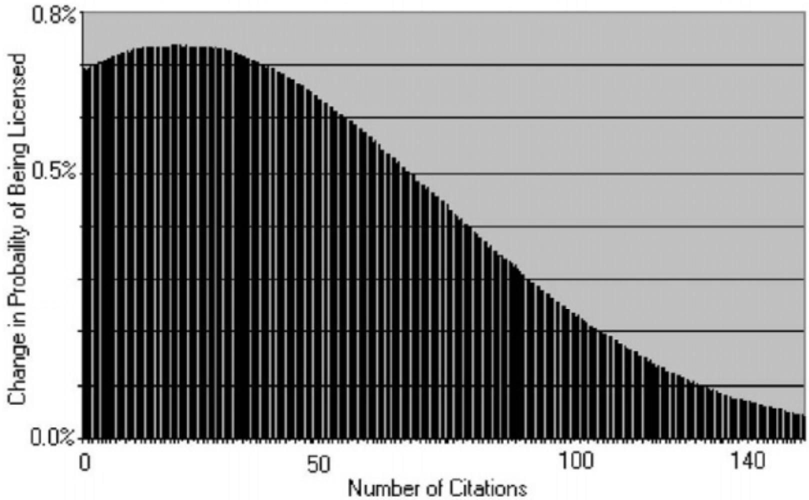


Figure 12.2: Predicted marginal effects, from Probit Model

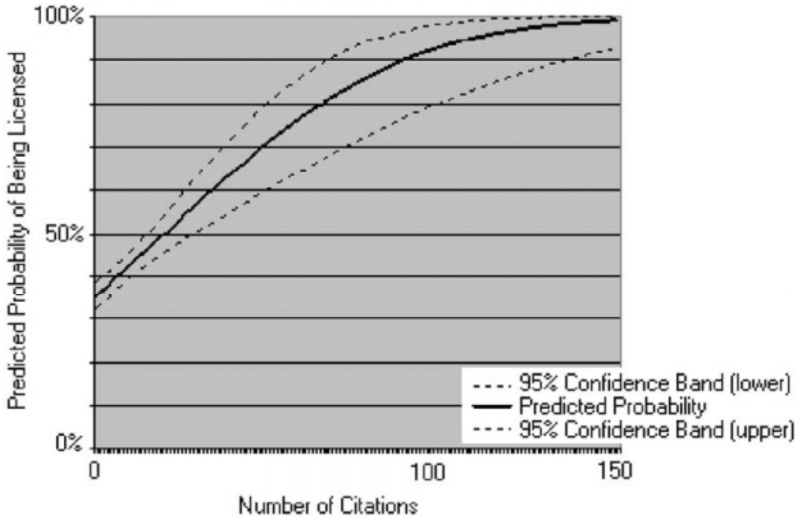


Figure 12.3: Predicted probability of being licensed, from Probit model

Figure 12.3 shows that as we approach the right tail of the citations distribution, the probability of licensing approaches 100%. The slope of the predicted probabilities is greatest for low values of citations (see also the marginal effects in Figure 12.2), suggesting that there is diminishing informational value to citations. To test this more formally, in separate regressions we include citations squared as an independent variable. The quadratic term is negative and statistically significant (at the 1% level), (consistent with diminishing informational content) but extremely small. The implied full marginal effect of citations does not change dramatically, except for extremely large values of citations. Consequently we exclude the quadratic term from the remaining specifications.

To check for possible differences in the citations–licensing relationship across universities, we estimate the baseline model for the University of California's patents alone and Columbia's patents alone. The results of these separate regressions are reported in Models 2 and 3 of Table 12.5. The probit coefficients for total citations are 0.017 and 0.029, respectively. Model 4 includes a term interacting citations with the Columbia dummy: the coefficient on this term thus gives differences across the universities in the effect of citations. The coefficient on the interaction term is small and statistically insignificant: we cannot therefore reject the hypothesis that the citations–licensing relationship is the same across the universities.

The probit results in Models 1–3 thus suggest a statistically and qualitatively significant relationship between citations and whether a patent is licensed. Citations provide information on whether an invention is licensed above and beyond what we can infer from time effects and technology class effects alone. These results are consistent with licensing policies practiced at Columbia and the University of California. Our interviews with licensing officers suggest that these two universities license all patents when it is possible to find a licensee. Moreover, these universities frequently require firms to underwrite patent application costs in exchange for a license. Thus the evidence so far suggests that citations are good indicators of whether there has been some commercial interest in a patent, consistent with Hypothesis 3.

Of course, licensed patents are only 'valuable' to a university in a sense of expected value. In order to examine whether citations reflect (or can predict) the actual returns from licensed inventions, we estimate tobit regressions analogous to the probit regressions above for the sub sample of patents assigned to Columbia University. Recall, as explained above, that in our analysis of licensing revenues, our unit of analysis is a licensed invention disclosure, or technology, rather than a patent. Tobit regressions are appropriate in this context because the dependent variable (log of $1 plus license revenues) is bounded below by zero. To be more precise, the source of the censoring is that amongst licensed technologies, for all observations for which a latent variable (such as 'quality') is not sufficiently high we shall observe zero revenues.

The results from the tobit regressions are presented in Table 12.6: Model 5 gives the basic results for Columbia University patents. For the overall sample the coefficients imply a marginal effect of 0.022 (i.e., a one unit change in citations leads to a 2.2% increase in revenues).[9] Although qualitatively significant, the effect is not statistically significant from zero in any of the models, however. These results suggest that citations do not appear to be good predictors of revenues earned by licensed technologies.[10]

Although citations are not related to license revenues in this sample, it is possible that citations by different types of citing entities are related to license revenues (as suggested by the hypotheses) and that their effect is being muddled by the other citations. To examine this possibility we disaggregated our sample by (a) pre-license citations versus post-license citations, and (b) by citations from licensees and citations from others. We estimate tobit regressions where we allowed these types of citations to enter separately (Models 6 and 7 of Table 12.6). As in the results for overall citations of these separate 'types' of citations have a statistically significant impact on revenues.

The lack of significance for coefficients of the licensee variables results suggest that citations are not related to license revenues, for licensed technologies, at least for the sub sample of patents licensed by Columbia University.[11]

[9] Marginal effects are calculated at the mean profile, using the procedure suggested by McDonald and Moffitt (1980).

[10] Though the tobit results provide little evidence of a relationship between citations and license revenues, these results may reflect the sensitivity of tobit specifications to violation of various statistical assumptions (Maddala, 1983). We therefore also ran but do not report non-parametric regressions analogous to the tobit model, but robust to violations of heteroskedasticity and non-normality assumptions, Powell's (1984) censored least absolute deviations (CLAD) estimator. CLAD regression results are qualitatively similar to the tobit results reported in Table 12.6.

[11] The results reported in Table 12.6 may reflect truncation bias if the citations–revenues relationship takes time to develop, however. Although these regressions control for time effects, the inclusion of later observations for which the relationship has not had sufficient time to develop might obscure any citations–revenue relationship in earlier observations. To check for this possibility we estimate, but do not report, tobit regressions where we include year dummies interacted with the citations term in addition to the application year dummies and the technology class dummies that are included in the Table 12.6 regressions. In this specification the coefficient on the interaction term gives the impact of citations within each application year. The coefficient of the interaction term is insignificant for most application years, again suggesting that citations are not good predictors of revenues for licensed technologies, and that the results above are not driven by truncation bias.

Table 12.6. Tobit Estimation for Licensed Technologies Only, Columbia University Patents

	Dependent Variable: Log(1 + License Revenues)		
	(5) Pooled Model	*(6) Pre- and Post- License Citations*	*(7) Licensee and Non-Licensee Citations*
Total Citations	0.024 (0.55)		
Licensee Citations		-0.440 (1.45)	
Non-Licensee Citations		0.059 (1.22)	
Post-License Citations			0.012 (0.24)
Pre-License Citations			0.096 (0.78)
1982 Dummy	-9.178 (1.22)	-34.173 (1.92)	-9.632 (1.28)
1983 Dummy	-5.810 (0.68)	-31.650 (1.69)	-6.275 (0.74)
1984 Dummy	-2.759 (0.37)	-28.621 (1.57)	-2.918 (0.39)
1985 Dummy	-8.685 (1.13)	-34.444 (1.88)	-9.170 (1.19)
1986 Dummy	-14.006 (1.61)	-39.870 (2.12)*	-14.503 (1.67)
1987 Dummy	-3.099 (0.37)	-28.507 (1.55)	-3.530 (0.42)
1988 Dummy	-8.053 (0.99)	-35.143 (1.82)	-8.857 (1.07)
1989 Dummy	-6.367 (0.73)	-32.806 (1.71)	-6.918 (0.79)
1990 Dummy	-7.974 (1.04)	-33.799 (1.84)	-8.391 (1.10)
1991 Dummy	-3.453 (0.40)	-28.929 (1.56)	-3.516 (0.41)
1992 Dummy	-5.735 (0.69)	-31.639 (1.70)	-6.142 (0.68)
1992 Dummy	-5.288 (0.63)	-30.883 (1.67)	-5.692 (0.74)
1994 Dummy	-3.199 (0.39)	-28.831 (1.56)	-3.514 (0.43)
Chemicals Dummy	-0.993 (0.16)	-2.441 (0.41)	-1.229 (0.20)
Drugs and Medical Dummy	-2.592 (0.44)	-3.873 (0.66)	-2.603 (0.44)
Electronics Dummy	-10.546 (1.80)	-11.868 (2.04)*	-10.764 (1.84)
Mechanical Dummy	-1.962 (0.24)	-1.995 (0.25)	-3.451 (0.41)
Constant	18.618 (1.98)	46.089 (2.30)*	19.160 (2.04)*
Number of Observations	56	56	56

5. DISCUSSION AND CONCLUSION

In the economics of innovation many important concepts are typically unobservable or difficult to measure systematically. Patent citation counts have been used as proxies for one such concept, the private value of patents. In this chapter, we have asked whether citations could predict two direct measures of value, whether a patent is licensed and how much revenue it generates conditional on licensing, for a sample of university patents where these measures were observable. Our primary finding is that whilst patent citations are good predictors of whether a university patent is licensed, they are not good predictors of the license revenues earned by technologies conditional upon its licensing.

How should we interpret the result that citations are good predictors of licensing, but not of revenues conditional upon licensing? One possibility, suggested by Hypothesis 3, is that citations reflect market interest in areas in technological proximity to particular patents. Market interest induces innovative effort in particular technological areas, increasing the probability of later citations. At the same time market interest also increases the probability of licensing. However, as innovation and commercialisation are uncertain activities, the level of revenues ultimately earned by particular technologies may be influenced by factors other than market interest, including competition by competing technologies, licensees' commercialisation incentives, and R&D and marketing competencies.

Thus we interpret the overall results as giving a preliminary nod to Hypothesis 3. The more fine grained tests of the mechanisms by which citations and private value are linked reveal little evidence in favour of Hypothesis 1 (citations reflect appropriability of spillovers), since licensees account for a small share of all citations, and licensing is the primary means through which universities could appropriate social returns. The occurrence of most citations after the license is executed provides some support for Hypothesis 2 (citations reflect entry into profitable areas) and Hypothesis 4 (citations reflect a disclosure effect), but that these later citations are not related to commercial success for the Columbia sub-sample suggests that these hypotheses do not explain the entire story. Clearly, more research remains to be done.[12]

[12] In assessing whether citations are useful proxies for private economic value, we have focused on the qualitative and statistical significance of the coefficients in regressions of value measures on citations. There may be other statistical dimensions upon which the usefulness of citations as proxies for value may be considered however. For example, from an errors in variables perspective, we would want to assure that the noise in citations (measurement error) is not related to the level of citations or the level of the underlying

We conclude by emphasising the preliminary nature of our results, particularly those relating to the citations–revenue relations. Since our tobit regressions rely on a small sample of technologies, the insignificant relationship between citations and revenues could be a result of sampling error. In future work we plan to use a much larger sample of technologies, including data from the University of California and possibly other universities.

More generally, prior research has highlighted the differences between university patents and those assigned to others, such as firms (Trajtenberg et al., 1997). Moreover, university motivations for patenting also vary (Mowery et al., 2001). Citations may therefore be more or less closely related to the 'value' of university patents than other patents — this is an open empirical question. If the reader prefers, she can consider this exercise a validation study of the relationship between citations and the value of university patents. The results presented here are material even under this more limited interpretation however, since an important use of citations in the economics literature has been as measures of the 'quality' of university patents (see Jaffe and Trajtenberg, 2002).

A decade ago Griliches (1990) noted that the use of patent citation measures "is only in its beginnings and we are likely to see a much wider use of it in the future" (p. 1689). As this statement remains true today, so does the need for more validation studies of the use of measures based on patent citations.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge support from Columbia University and the Mack Center for Technological Innovation at the Wharton School of the University of Pennsylvania. We thank Rebecca Henderson, David Mowery, Richard Nelson, and participants at the NBER Productivity Meeting for thoughtful comments on earlier versions of this research.

---

variable of interest (see Krasker and Pratt, 1986). Such an analysis remains for future work.

REFERENCES

Association of University Technology Managers 2000. Licensing Survey, FY 1999.

Austin, David H. (1993). An event–study approach to measuring innovative output: the case of biotechnology. *American Economic Review,* 83, 253–258.

Austin, David H. (1994). *Patent citations and appropriability*. Resources for the Future.

Buderi, Robert (1999). In search of innovation. *Technology Review*.

Caballero, Ricardo J., Jaffe, Adam B. (1993). How high are the Giants' shoulders: an empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. MIT and NBER; Harvard University and NBER.

Carpenter, M., Narin, F., Woolf, P. (1981). Citation rates to technologically important patents. *World Patent Information*, 4, 160–163.

Caudill, S., Jackson, J. (1989). Measuring marginal effects in limited dependent variable models. *The Statistician,* 38, 203–206.

Cole, Jonathan R. (2000). *A short history of the use of citations as a measure of the impact of scientific and scholarly work*. In Blaise Cronin, Helen Atkins (Eds.), The Web of Knowledge: A Festschrift in honor of Eugene Garfield (pp. 281–300. Medford, N.J.: Information Today.

Deng, Zhen, Lev Baruch, Narin, Francis (1999). Science & technology as predictors of Stock performance. *Financial Analysts Journal,* 55.

Directory of Corporate Affiliations, 1983–2000. Skokie, Ill.: National Register Pub. Co.

Griliches, Zvi (1986). *Economic data issues*. In Zvi Griliches, Michael D. Intriligator (Eds.), Handbook of econometrics (pp.1466–1514). Volume 3. Handbooks in economics series, Book 2. Amsterdam, Oxford and Tokyo: North-Holland

Griliches, Zvi (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature,* 28, 1661–1707.

Griliches, Zvi (1992). The search for R&D spillovers. *Scandinavian Journal of Economics,* 94, S29–47.

Hall, Bronwyn H., Jaffe, Adam, Trajtenberg, Manuel (2000). *Market value and patent citations: A first look*. NBER Working Paper W7741.

Harhoff, Dietmar, et al. (1999). Citation frequency and the value of patented inventions. *Review of Economics and Statistics,* 81, 511–515.

Harhoff, Dietmer, Scherer, Frederic, Vopel, Katrin (1999). Citations, family size, and the value of patent rights. *Mimeo*.

Henderson, Rebecca, Jaffe, Adam B., Trajtenberg, Manuel (1998). Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988. *Review of Economics and Statistics,* 80, 119–127.

Jaffe, Adam (1998). Patents, patent citations, and the dynamics of technological change. *NBER Reporter,* 1998.

Jaffe, Adam, Henderson, Rebecca, Trajtenberg, Manuel (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics,* 108.

Jaffe, Adam, Trajtenberg, Manuel, Fogarty, Michael (2000). The meaning of patent citations: report on the NBER/Case-western reserve survey of patentees. *NBER Working Paper W7631*.

Jaffe, Adam B., Fogarty, Michael S., Banks, Bruce A. (1998). Evidence from patents and patent citations on the impact of NASA and other federal labs on commercial innovation. *Journal of Industrial Economics,* 46, 183–205.

Jaffe, Adam B., Trajtenberg, Manuel (1999). International knowledge flows: evidence from patent citations. *Economics of Innovation and New Technology,* 8, 105–136.

Jaffe, Adam B., Trajtenberg, Manuel (2000). *Patents, citations & innovations: A window on the knowledge economy.* (Cambridge, MA: MIT Press).

Jensen, Richard, Thursby, Marie (1998). Proofs and prototypes for sale: The tale of university licensing. *American Economic Review* .

Klepper, Steven, Leamer, Edward E. (1984). Consistent sets of estimates for regressions with errors in all variables. *Econometrica,* 52, 163–183.

Krasker, William S., Pratt, John W. (1986). Bounding the effects of proxy variables on regression coefficients. *Econometrica,* 54, 641–655.

Krugman, Paul (1991). *Geography and Trade*. Gaston Eyskens lecture series, Cambridge, Mass. and London.

Lanjouw, Jean O., Schankerman, Mark (1999). The quality of ideas: measuring innovation with multiple indicators. *NBER Working Paper W7345.*

Lanjouw, Jean O., Schankerman, Mark (2000). *Characteristics of patent litigation: a window on competition*. Draft Manuscript.

Maddala, G.S. (1983). *Limited–dependent and qualitative variables in econometrics*. New York: Cambridge University Press.

McDonald, John F., Moffitt, Robert A. (1980). The uses of tobit analysis. *Review of Economics and Statistics,* 62, 318–321.

Melkers, Julia. (1993). *Bibliometrics as a tool for analysis of R&D impacts*. In Barry Bozeman, Julia Melkers (Eds.), Evaluating R&D impacts: Methods and practice. Boston: Kluwer.

Mogee, Mary Ellen, Kolar, Richard G., Putnam, Jonathan D. (1997). *Patent indicators of global technology strategy*. Reston, VA: Mogee Research & Analysis Associates.

Mowery, David C., Sampat, Bhaven N. (2001). Patenting and licensing university inventions: lessons from the history of research corporation. *Industrial and Corporate Change*.

Mowery, David C., Sampat, Bhaven N. (2001). University patents and patent policies: 1925–1980. *Industrial and Corporate Change*.

Mowery, David C., Sampat, Bhaven N., Nelson, Richard R., Ziedonis, Arvids A. (2001). The growth of patenting and licensing by U.S. universities: an assessment of the effects of the Bayh–Dole act of 1980. *Research Policy,* 30, 99–119.

Narin, F., Noma, E., Perry, R. (1987). Patents as indicators of corporate technological strength. *Research Policy*, 16, 143–155.

Powell, James L. (1986). Censored regression quantiles. *Journal of Econometrics,* 32, 143–155.

Scherer, F.M., Harhoff, Dietmer (2000). Technology policy for a world of skew-distributed outcomes. *Research Policy,* 29, 559–566.

Shane, Hilary, Klock, Mark (1997). The relation between patent citations and Tobin's Q in the semiconductor industry. *Review of Quantitative Finance and Accounting,* 9, 131–146.

Trajtenberg, Manuel (1990). A penny for your quotes: patent citations and the value of innovations. *Rand Journal of Economics,* 21, 172–187.

Trajtenberg, Manuel, Henderson, Rebecca, Jaffe, Adam (1997). University versus corporate patents: a window on the basicness of invention. *Economics of Innovation and New Technology,* 5, 19–50.

Chapter 13

SCIENTIFIC AND TECHNOLOGICAL PERFORMANCE BY GENDER

Fulvio Naldi[1,3], Daniela Luzi[2], Adriana Valente[2], and Ilaria Vannini Parenti[3]

[1]Consiglio Nazionale delle Ricerche , Milano, Italy. E-mail:naldi@area.mi.cnr.it
[2]Consiglio Nazionale delle Ricerche-Istituto di ricerche sulla popolazione e le politiche sociali. Rome, Italy. [3]DNSistemi srl, Milano, Italy

Abstract: The availability of sex-disaggregated data in the fields of research, technology and development is extremely important for supporting the growing political commitment to promote and monitor women participation in the different fields of S&T. During the late 1990s the European Commission identified as a priority the availability of this data. Even if scientific publications and patents are widely accepted indicators of scientific and technological performances, until now it has been impossible to measure bibliometric and patent output by gender in a large set of data. Starting from a feasibility study carried out for the European Commission on the whole set of patents published in 1998 by the European Patent Office and on 30,000 authors of items published in 1995 on scientific journals of international relevance, the paper demonstrates that it is possible to obtain robust gender indicators on S&T output.

1. INTRODUCTION

Many metaphors are generally used in gender studies on science and technology (S&T): *leaky pipeline, crystal glass ceiling, scissors effect, impossible pursuit, overtaking model* (Palomba, 2000). These metaphors give both the idea of an inefficient use of human resources and of invisible constraints which bias scientific performance. They refer to the difficulties of women in reaching career levels comparable to those of their male colleagues, merits and education levels being equal. This waste of resources takes place in different ways and at different times of women's involvement in science.

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 299-314.

At the beginning young female students are not encouraged to undertake full time research professions and therefore "they are lost to the scientific world" (Palomba, 2000) like a *leaky pipeline* which continues through their entire career. Invisible obstacles, described as *a crystal glass ceiling*, prevent women from career progression. The *scissors effect*, which represents the most stable, statistically measurable phenomenon, indicates the progressive split between male and female careers. The two metaphors *impossible pursuit* and *overtaking model* refer to female presence in the different scientific fields: women cannot overcome the initial disadvantage of being under-represented in *hard sciences*, so that their pursuit is impossible. In sectors in which women are highly represented at the beginning of their career, they become a minority whilst proceeding to the top positions; one possible explanation is that in these sectors it is more difficult to find attractive professional positions outside academies and research institutions.

The growing importance of the issue of gender presence in science and technology led to defining indicators which could explain inequalities and differences between groups in terms of vertical and horizontal segregation (Siltanen et al. 1995). The former indicates the share of women in research activities and/or in specific disciplinary fields, and it has the implicit assumption that a more balanced distribution of women may be a good result in terms of gender equalities. The vertical differences, although strictly related to the horizontal differences, investigate the distribution of women throughout the scientific career ladders. Close to the issue of vertical segregation the reward and recognition system of European Universities and scientific institutions has been analysed (Harding et McGregor, 1996; Osborn, 2000).

The statistical description of the participation of women in S&T sectors is developing step by step. Until now the efforts for collecting gender data in Europe has not already produced harmonised and comparable sex-desegregated data for R&D and for S&T human resources. Problems of homogeneity and completeness of data still have not been solved. As a matter of fact, 'general purpose' data sets, which include demographic data and labour force surveys, may only be used for basic analysis, whilst dedicated surveys tend to lack coverage and representativeness.

There is a need for promoting gendering and statistical collection at national and European level (the top–down approach), as repeatedly recommended by the European Commission (European Commission, 2003) which considers the collection of data on scientific publications distributed by gender as a long term important task to achieve. In the meantime the collection of existing data at national level (the bottom–up approach) may show particular/local contexts and produce an insight into the development of new gender indicators.

This paper summarises a project (Naldi and Vannini Parenti, 2002) carried out for the European Commission – DG Research, aimed at assessing the feasibility of producing patent and bibliometric indicators by the gender of the inventor/author. It would appear to be the first study of its kind, and the results provide some pioneering measures of sex desegregated S&T output and productivity.

2. ANALYSES OF SCIENTIFIC PERFORMANCE AND POLICY FOR GENDER MAINSTREAMING

The wastage of women's skills and knowledge weigh heavily in the science system. First gender studies on science have posed the question of male and female scientific productivity in the US; they showed that female scientists produce a lower number of publications, and are less cited (Cole and Cole, 1973). Some studies in Europe followed a comparable approach and reached the same conclusions (Bochow and Joas, 1987). These results have been explained only in terms of family roles and workloads, but these variables have been considered overestimated by further studies (Zuckerman, 1992). All studies cited from the European Report on Science and Technology Indicators of 2003 show, on the contrary, a relationship between familiar factors and women's careers, even if each one uses different methodologies.

In the 1990s studies at EU level and inside the Member States showed the complexity of the phenomenon that cannot be easily explained without taking into account different perspectives and variables related to the context of S&T. Many factors may influence scientific productivity: the structure and organisation of scientific communities, the selection criteria of accessing postgraduate education and professions, the evaluation procedures of applications and grants of research funding as well as the participation to commissions and evaluation committees.

Analyses carried out in both Europe and in the US showed a close relationship between the scientific production and career levels (Long, 1992; Kaplan et al., 1996). Other studies stressed, in particular, the quality of publications (Campanelli et al., 1999) and citation patterns (Sonnert and Holten, 1996). Two studies (Long, 2001; Di Cesare, Luzi, Valente, 2003) have analysed the relationship between careers and publications. Long demonstrated that the male full professors of universities publish 30% more than their female colleagues. The second analysis related to male and female researchers of the Italian National Research Council extracted from the Social *Science Citation Index* for 1999–2001, demonstrated the increased attitude of male and female researchers to publish when belonging to a high grade hierarchy.

The lack of longitudinal studies represents a limit to interpretation: Xie and Shauman (1998) showed a decrease in difference of scientific production considering national inquires in the period of time 1969 to 1993.

The difference of scientific production depends partially from the overestimation of males in extremely productive groups and the lack of consideration of the part-time jobs of female researchers. (European Commission, 2003B)

At the end of the 1990s the European Commission promoted both specific and long term actions; the first type of action was to commission a report on women and science in the EU to an *ad hoc* group, the European Technology Assessment Network (ETAN) group. The report, published in 2000, in dealing with gender inequality in science, highlighted the phenomenon of the 'leaky pipeline'. From the ETAN report forth, this phenomenon has been further confirmed without distinction of countries and disciplines; even countries with advanced equality legislation are experiencing the situation and consequences of the *leaky pipeline*. This phenomenon represents one of the cases in which, besides quantitative studies and integrated data, qualitative analyses become crucial for a better understanding of the causes of this discrimination and waste of resource as well as for the identification of positive actions. The ETAN group also recommended that Women and Science Units would be present in all State Members, and this recommendation has been adopted by many Member States.

As a long term policy the Commission set out an action plan to promote gender equality in science and appointed a group of experts (known as the Helsinki Group) which meet on a regular basis. The group guarantees exchanges of experiences on measures and policies introduced in different countries and provides sex desegregated statistics, thus allowing continuous monitoring and promotion of the participation of women in S&T. One of the most important achievements of the group has been the delivering of national reports, based on a common structure, in which the collection of information on policy measures are as relevant as statistical data on women in science.

The focus on national policies was one of the commitments for the Member States identified by the Research Council as a priority, together with information exchange about human resources in S&T and common procedures of collecting data. Data collected by the Helsinki Group show that the organisation of the scientific system is quite similar among member states, being most of the responsibility centred in Ministries often devoted also to education and cultural issues. Member states also have weak points in common: under representation of women and lack of gender balance, on the top career ladders and decision levels.

Differences arise with reference to policy context and to the different measures to promote gender equality: the presence of a Ministry for Women or of a Women and Science Unit inside the Science Ministry; the provision of targets and quotas for a gender balance on university/research institute committees; the development of gender equality indicators; the production by Universities and Research Institutes of equality plans. Besides positive actions like supporting networks of women in science and establishing targets, quotas, research funds and prices for women, some countries have recently been considering gender mainstreaming, which is the systematic integration of gender equality into all policies and programmes and is embedded into EU policy.

Gender mainstreaming measures focus on 'legislation' and 'gender studies'. The latter is a strategic issue to increase the knowledge for better comprehension and interpretation of the phenomenon, as institutional practices seem to produce, albeit unintentionally, discriminatory effects that cannot be changed by new legislation alone. 'Gender proofing pedagogy of science education', relates to gender differences in the methods and content of teaching and involves the question of values of science. 'Work/life balance measures' which includes part-time as well as time flexibility, are examples of tools which may benefit men as well as women. Another mainstreaming tool is 'modernising human resource management', that includes transparency in appointment, promotion and recruitment procedures, as well as the reinterpretation of the concepts of merit and excellence. This is a very challenging task, considering that in the 1980s Merton had already pointed out that 'reward' and 'excellence' have instrumental and honorific, thus not objective, meaning, which makes evaluation activity very difficult.

3. FIRST NAMES AS A TOOL FOR GENDER CLASSIFICATION

Patent and bibliographic databases do not contain coding on gender of inventors and authors. To overcome the problem a feasibility study has been performed to verify the effectiveness of genderise data on patents and scientific publications by using the first names of authors and inventors.

For this purpose a comprehensive 'First Name Data Base' (FNDB) was created and applied to a significant sample of patents and scientific publications. The current release of FNDB covers 6 European languages: English, French, German, Italian, Spanish, and Swedish and contains 8,291 different names selected from more than 32,000 names collected from different sources such as dictionaries, calendars, books and internet sites, files from Record Offices, and phone books. In FNDB, 3,634 names are classified as female, 4,115 as male, and 452 are commonly used for both genders or are language dependent.

The setting up of a high quality database had two objectives: (1) to perform gender analyses on any list of first names, and (2) to allow expansion to other languages. Each name is classified by gender, following a classification which is language/country dependent, to solve cases in which the same name belongs to different genders in different languages. This is the case for example of ‘Andrea’ which is male in Italian and female in Spanish and German. The adopted strategy improves data quality and reliability and is described in details in the Final Report of the project (Naldi and Vannini Parenti, 2002) together with the techniques developed to manage diacritics, double names, exceptions, etc.

The degree of coverage of FNDB has been tested on more than 100,000 names of inventors and on about 30,000 names of authors of scientific papers. The results are summarised in Table 13.1.

Table 13.1. DB coverage by country/language

	Inventors					*Authors*				
Country	Total	Not found	%	Both	%	Total	Not found	%	Both	%
DE	55,195	842	1.5	89	0.2	6,865	257	3.7	51	0.7
ES	1,383	44	3.2	12	0.9	2,766	166	6.0	62	2.2
FR	16,973	239	1.4	524	3.1	6,030	191	3.2	228	3.8
GB	15,979	420	2.6	197	1.2	7,468	487	6.5	237	3.2
IT	6,745	106	1.6	12	0.2	5,202	104	2.0	18	0.3
SE	6,718	296	4.4	56	0.8	1,528	114	7.5	25	1.6
Total	102,993	1947	1.9	890	0.9	29,859	1,319	4.4	621	2.1

The adopted methodology is successful in more than 90% of cases: 97.2% of the inventors and 93.5% of authors were identified by FNDB. The unidentified inventors and authors remained unclassified because their names were not included in the database (1.9% and 4.4%) or are currently used for both genders (0.9% and 2.1%). Coverage of patents is strongly influenced by German inventors who represent more than 50% of the total number of the names. The sample of authors of scientific publications is better distributed amongst the 6 Countries but contains a larger number of ‘foreign’ people (mainly from Arabic and Far Eastern countries) who are working in the 6 countries and whose names are not included in FNDB.

A further demonstration of the feasibility of the methodology comes from the analysis of the distribution of missing names by number of occurrences: 73% inventors and 90% authors of the missing names appear only once in the database. These names can be considered as spelling errors and rare, or foreign, names.

4. GENDER ANALYSIS ON R&D OUTPUT

This study has been performed on two sets of data:

- Patents published in the year 1998 by the *European Patent Office* (EPO) and produced by inventors whose working address is in France, Germany, Italy, Spain, Sweden, and the UK.
- Scientific papers published in the year 1995 in 157 scientific journals of international relevance by authors of the same 6 EU countries.

4.1 Notes on the Adopted Methodology

The EPO database already includes the first names of the inventors. For this reason it has been possible to process the whole set of 47,820 patents and 102,993 inventors published for the year 1998 from the 6 countries.

Patents are classified according to the International Patent Classification (IPC) Schema. Up to 4 IPC codes are assigned to each patent. Correspondence tables have been applied to assign patents to Industry Sectors (Verspagen, Moergastel, Slabbers, 1994) and Field of Technology. The same patent can be assigned to more than one Industry Sector / Field of Technology.

Differently from the patents databases, bibliographic databases do not contain the first names of the authors but only their initials. For this reason it was necessary to collect the names manually from the original paper, where, however, the first name is available in less than 50% of cases. The sampling procedure was based on an '*a priori*' selection of the journals. Journals were selected on the basis of the high availability of authors' first names, high frequency of items written by authors of one of the 6 countries, high scientific relevance, and balance of the geographical and disciplinary coverage. Since it was impossible to predict in advance the number of first names actually available in the chosen journals, the sample was built in a dynamic way, carrying out adjustments in real time during data collection, with the goal of collecting a significant number of authors for each country and discipline. Extra data have been collected for Medicine, Chemistry and Physics in order to check the sampling methodology and suggest possible future extension of the analysis. A sample of 9,344 publications and 36,239 authors was obtained after processing more than 100,000 items published in the selected journals.

Each publication was classified according to the disciplinary sector(s) of the journal in which it was published. The *Science Citation Index* (SCI) '95 classification of the journals, based on 183 disciplines, was used. SCI disciplines have been grouped into 9 disciplinary sectors: Biology (Biol), Biomedical Research (Biomed), Chemistry (Chem), Clinical Medicine (Med), Earth and Space (Earth), Engineering (Eng), Mathematics (Math), Physics (Phys), Multidisciplinary Sciences (Mult). Some disciplines may be associated with more than one sector. Journals may be associated with more than one discipline and may belong to more than one sector.

Three indicators were introduced in order to evaluate patents and publications produced by co-operation among inventors/authors of different countries and gender:

– *Participation* counts the number of patents/publications with at least one author of a given gender/country;
– *Contribution* measures the involvement of each gender/country in the production of a patent/publication, assuming that each person contributed the same amount. *Contribution* is also called *'patents/publication-equivalent'* since it sums up the single shares of each item attributed to a given gender/country. In general, for a patent/publication with *n* authors the contribution of each gender/country is equal to the number of authors of the respective gender/country divided by *n*. The sum of the *contributions* of all the genders/countries involved in a patent/publication is always equal to 1;
– *Presence*: Total count of the authors of a given gender/country in each patent/publication.

4.2 Distribution by Gender

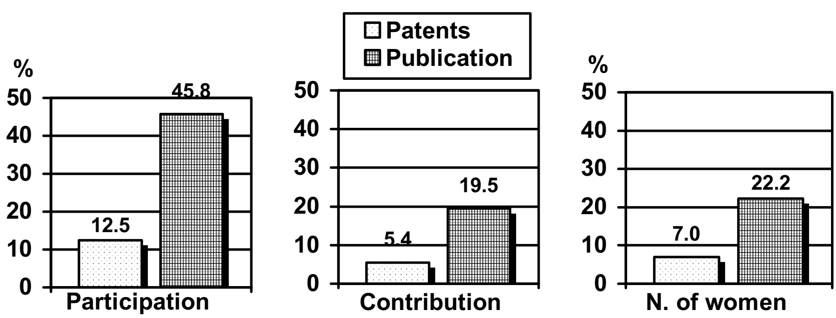


Figure 13.1. Participation, contribution and presence of women in patents and publications

Figure 13.1 shows the participation, contribution, and presence of women respect to the total number of inventors/authors.

The patents with at least one female inventor are 12.5% whilst 97.3% of the patents have at least one male inventor. As a consequence 87.5% of the patents has been entirely produced by men and 2.7% entirely by women. On the other hand female inventors are 7% of the total number and contribute to the overall production of patents with 5% of equivalent patents. It is important to note that since one half of the patents are produced in Germany, the low percentage of German female inventors influences significantly the global statistics.

The publications with at least one female author are 45.8% whilst the items with at least one male author are 94.7%. As a consequence 54.2% of the items have been entirely produced by men and 5.3% entirely by women. On the other hand female authors constitute 22% of the total and contribute 20% of equivalent publications to the overall scientific production.

4.3 Distribution by Country

Figure 13.2 shows the contribution of women to patents and publications in the 6 countries. In analysing the statistics on patents, the geographical bias has to be taken in consideration: German inventors are almost one half (48%) of the total and are involved in 44% of the patents. French and British inventors represent both 15% of the total, Italy and Sweden about 6%, Spain only 1.2%.

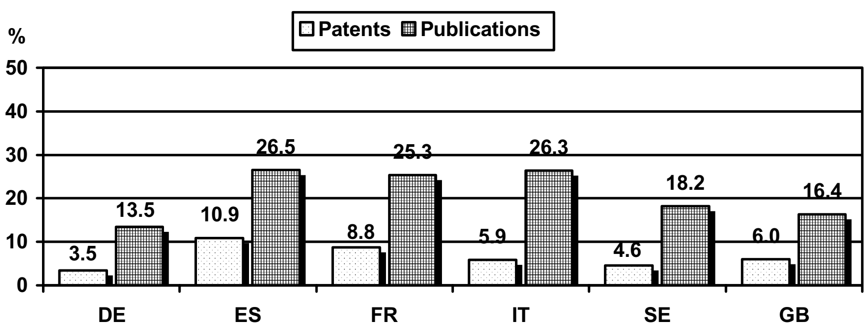


Figure 13.2. Female contribution by country to patents and publications

The country with a higher female contribution to patents is Spain followed by France. Scientific publications show two patterns of countries

— Italy, Spain and France with a relatively high female contribution and Germany, Britain and Sweden with significantly lower contribution.

The relevance of the differences in gender distribution amongst countries can be pointed out by observing that the percentage of German females is nearly half of those of Spain, Italy and France and that, for example, the total number (both men and women) of *publication-equivalents* produced by UK is more than twice that of Italy (2,387 compared with 1,121) but the total number of British female authors (1,260) is lower than the number of Italian female authors (1,426).

The statistics on participation confirm the trends of contribution shown in Figure 13.2. Spain and France have the highest percentage of patents with at least one female inventor (19.4% and 16.8% respectively) and Germany has the lowest percentage of female inventors (4.6% vs 15.8% of Spain). The percentages of publication with at least one female author in Italy, Spain and France (respectively 58.3%, 56.3% and 53.6%) is remarkably higher than that of Sweden (38.0%), United Kingdom (31.8%) and Germany (32.3%).

It can be noted that Sweden, which has a long tradition and practice in supporting gender policy, is just above the United Kingdom.

These results look less surprising if we compare our data with those provided by the WIS database of the European Community (European Commission, 2003B) (Figure 13.3): the percentage of female authors looks related to the share of global female labour force in the public sector (government and higher education sector) in the six countries considered. Public female researchers in Spain have the highest share, followed by Italy. France and Sweden are at the same level, whilst the United Kingdom and Germany are in the last position.

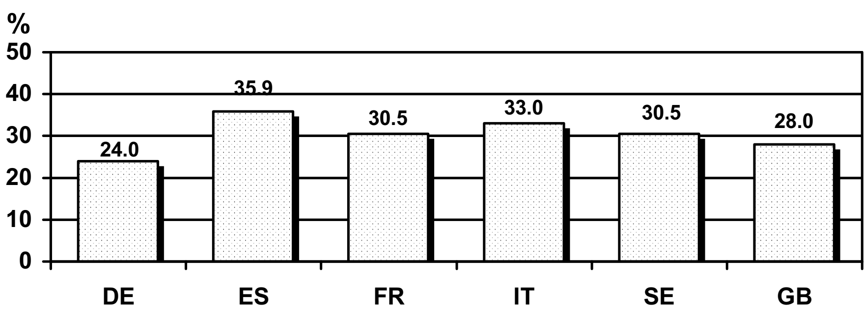


Figure 13.3. Share of women researchers in the two public sectors (GOV, HES). Source: elaboration from WIS database, DG Research

Even if the data of Figure 13.3 refer to the year 1999 and only focus on researchers of the public sector, they may give interesting clues for interpretation and further analyses. We have to consider, for instance, that countries such as Sweden, the United Kingdom, and Germany have a high percentage of researchers working in the private sector; anyway this should not significantly influence the number of publications, because the private sector tends not to publish as much as the public sector.

4.4 Distribution of Patents by Industry Sectors

The following figure (13.4) shows the number of equivalent patents produced in each sector by female inventors, expressed as percentage of the total number of patents in the sector.

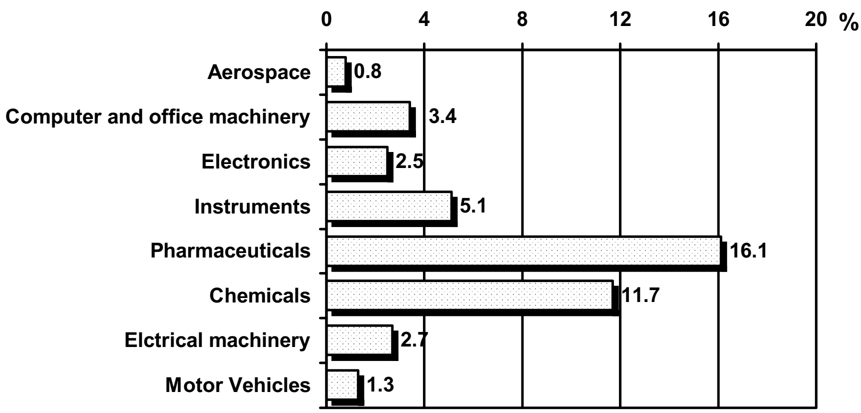


Figure 13.4. Female contribution to patents by Industry sectors

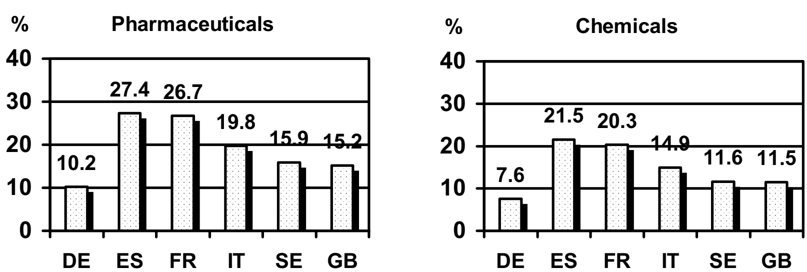


Figure 13.5. Female contribution by country to patents in Pharmaceuticals and Chemicals

Distribution by country of the two sectors where the female contribution exceeds 10% is reported in the following Figures 13.5 and 13.6.

As a general consideration Germany confirms its position of leadership in all the sectors for the general ranking, maintaining the first place both for number of patents and for total inventors. On the other hand, Germany has the lowest percentage of women in almost all the sectors whilst France and Spain have a strong presence of women in most fields.

4.5 Distribution of Scientific Publications by Discipline

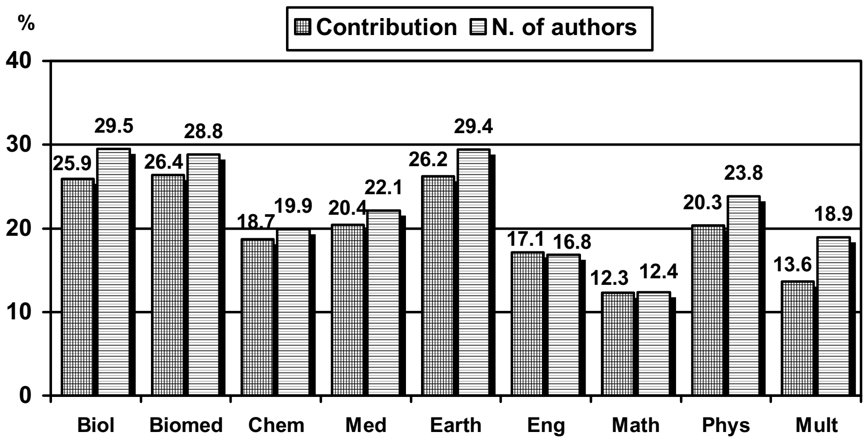


Figure 13.6. Female contribution to scientific publications by discipline

Whilst the percentage of female authors is significantly above the average in Biology, Biomedicine, and Earth and Space Sciences, it is below the average in Engineering, in the Multidisciplinary sector and in Mathematics ($\chi^2 = 335{,}991$, df = 8, $p < 0.001$).

Some peculiarities which arose from the statistics can be pointed out and deserve further analysis:

- The *participation* of women in Mathematics is remarkably low. This is only partially justified by a general (*i.e.* independent of gender) low level of co-authoring in this discipline.
- Engineering is the only discipline in which the *contribution* of women is greater than the percentage of the *number of authors*.
- Clinical Medicine is the discipline where the difference between the two groups of countries is less evident although still significant.
- Italian female authors in Biomedicine participate in about 80% of the publications with more than 40% of article equivalents.

– The generally small presence of women in Mathematics is particularly low (well below 10%) in Germany, Britain, and Sweden.
– There is a very high percentage of Swedish women in Earth and Space Sciences. This data should be further analysed with a larger data sample.

4.6 Other Bibliometric Indicators

4.6.1 Gender by type of publication

The cross-tabulation between gender and type of publication shows that the percentage of female authors in the types of publication traditionally used to communicate scientific results: articles (22.7%); letters (21.9%); and notes (21.4%); is significantly higher ($\chi^2 = 63{,}052$, df = 5, $p < 0.001$) than in publications relating to editorial activity: editorials (10.6%) and reviews (14.8%). This can be explained either by a lower level of interest of females in the editorial activity of the journals (editorial, notes, etc.) or by some kind of discrimination in the editorial management.

4.6.2 Distribution by first authors

This analysis was carried out on the 6,159 items of the sample with two or more authors and where the authors were not listed in alphabetical order. No significant differences were found between the gender distribution of first authors and the gender distribution of all the authors of this specific sample

4.6.3 Single authors

This analysis was carried out on the 1,570 items written by single authors. In the publications with only one author the female contribution is 10.8%, significantly ($\chi^2 = 113{,}983$, df = 1, $p < 0.001$) smaller than the whole data sample. Figure 13.7 shows the female contribution of single authors by country.

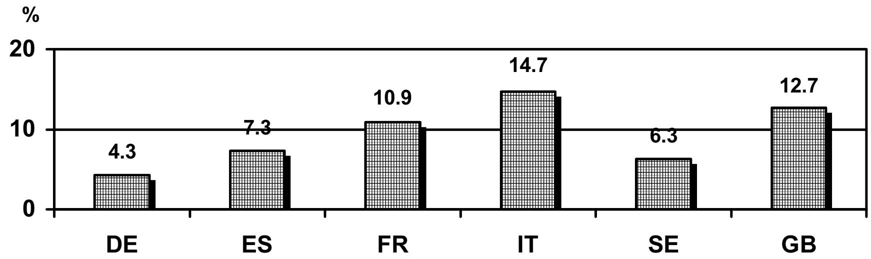


Figure 13.6. Percentage of women among first authors

4.6.4 Co-authoring

Figure 13.8 shows the contribution of women as a function of the number of co-authors. The contribution of women, and not only their participation, increases with the number of co-authors. That could indicate a better inclination of women to co-operate and to participate in large research groups.

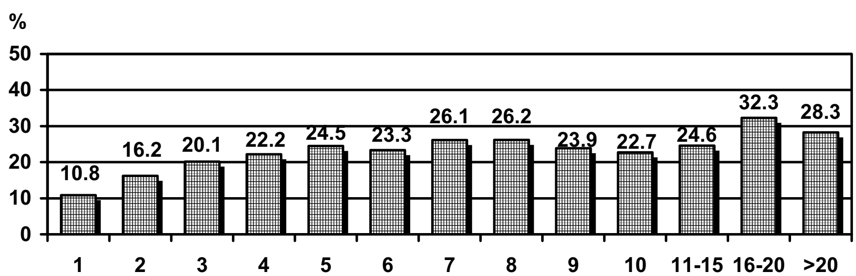


Figure 13.8. Female contribution by number of co-authors

5. CONCLUSIONS

In this paper we confirm the feasibility of the approach of using first names to produce robust gender indicators which can be applied to any data set containing first names. The statistical tables presented are highly reliable for the size of the sample analysed.

As a further step, in our opinion, the study should be extended to cover more countries and a larger period of time in order to include at least all EU countries and to evaluate the trends also in connection with the political actions promoted at national and international level.

More particularly, even if the mean contribution of female inventors to patents is still relatively low, there are some technologies in which the participation of women is highly significant. For these fields, an accurate analysis of temporal data would provide important indicators on the presence of women in the sector of industrial R&D, a situation which up to now has not been explored in detail.

As for the interpretation of the statistics on scientific productivity, it is worthwhile considering the different policies of publication and the chosen channels of dissemination in the different disciplines. Even if the international journals are favourite, in some cases (e.g. social sciences and humanities) the authors' preference goes to monograph publications, which are excluded from the citation indexes. Depending on the disciplinary sectors, some parameters, such as the number of authors per publication and the yearly mean scientific production, may vary, as well as the number of journals included in the principal citation indexes. For this reason an exhaustive investigation should take a broader set of sources into account and include social sciences, the arts and the humanities. Moreover, some scientific communities, such as physics, mathematics, computer science, start giving great importance to the diffusion of results trough Open Archives. With the increasing prestige and number of publications available on Open Archives gender analyses have also to take these new channels of diffusion into account. They can also turn out to be an important tool for collecting authors' first names more easily, facilitating the bibliometric analysis of publications.

In the future it would be useful to connect data on scientific productivity with other variables, such as the number and position of female and male scientists and researchers, which can provide a new perspective for analysing more deeply the question of gender in scientific performance. Moreover, it would be necessary to introduce objective measures on the way of working in the scientific world (Palomba, 2000) in addition of the reinforcement/improvement of qualitative and quantitative studies.

REFERENCES

Bochow, M., Joas, H. (1987). *Wissenschaft und Karriere, der berufliche Verbleib der akademischen Mittelbaus.* Frankfurt am Main: Campus.

Campanelli, G., Segnana, M.L., Soci, A. (1999). *Attività didattica, visibilità e pubblicazioni dei giovani economisti italiani. Una prospettiva di genere*. In Carabelli, A., Parisi, D., Rosselli, A., 1999, Che genere di economista? Bologna: Il Mulino.

Cole, J.R., Cole, S. (1973). *Social stratification in science*. Chicago: University Chicago Press.

Di Cesare, R., Luzi, D., Valente, A. (2003). *La produzione scientifica del Cnr nelle scienze sociali: considerazioni di genere*. In A. Valente, D. Luzi (Eds.), Partecipare la scienza. Roma: Biblink (in press)

European Commission — Directorate-General for Research, Science and Society. (2003). *She figures: women and science, statistics and indicators*. EUR 20733.

European Commission — Directorate-General for Research, Knowledge–based economy and society competitiveness, economic analysis and indicators (2003). *Third European Report on science and technology indicators, towards a knowledge based economy*. Office for Official Publications of the European Communities, Luxembourg.

Harding, S., Mc Gregor, E. (1996). *The conceptual framework*. In World Science Report (pp. 303–324). Paris: Unesco.

Kaplan, S.H., Sullivan, L.M., Dukes, K.A., Phillis, C.F., Kelch, R.P., Schaller, J.G. (1996). Sex differences in academic advancement, *England Journal of Medicine*, 335, 1282–1289.

Long, J.S. (Ed.). (2001). *From Scarcity to Visibility: Gender Differences in the Careers of Doctoral Scientist and Engineers*. Panel for the study of gender differences in career outcomes of science and engineering Ph.D.s. Washington: National Academy Press.

Long, J.S. (1992). Measures of sex differences in scientific productivity. *Social Forces*, 71 (1), 159–178.

Naldi, F., Vannini Parenti, I. (2002). *Scientific and technological performance by gender — A feasibility study on patents and bibliometric indicators*. Luxembourg: Office for Official Publications of the European Communities.

Osborn, et al. (2000). *ETAN Report science policies in the European Union – promoting excellence through mainstreaming gender equality*. Luxembourg: Office for Official Publications of the European Communities.

Palomba, R. (ed.) (2000). *Figlie di Minerva*. Milano: Franco Angeli.

Siltanen, J., et al. (1995). *Gender inequicity in the labour market*. Geneva: ILO.

Sonnert, G., Holton, G. (1996). Career pattern of women and men in sciences. *American Scientist,* 84, 67–71.

Verspagen, B., van Moergastel, T., Slabbers, M. (1994). *MERIT concordance table: IPC - ISIC (rev.2)* . RM1994-004, http://www.merit.unimaas.nl/publications.

Xie, Y., Shauman, K.A. (1998). Sex differences in research productivity. New evidence about an old puzzle. *American Sociological Review, Official Journal of the American Sociological Association*, 63, 847–870.

Zuckerman, H., Cole, J.R., Bruer, J.T. (Eds.). (1992). *The outer circle, women in scientific community*. New Haven and London: Yale University Press.

Chapter 14

THE USE OF INPUT DATA IN THE PERFORMANCE ANALYSIS OF R&D SYSTEMS

Potentialities and Pitfalls

Marc Luwel

Centre for Science and Technology Studies, University Leiden, the Netherlands; and Dutch-Flemish Accreditation Organisation, The Hague, the Netherlands
E-mail: luwel@cwts.leidenuniv.nl; m.luwel@nvao.net.

Abstract: With the emergence of the knowledge based society, great emphasis is put on the development of qualitative and quantitative policy tools for analysing the science and innovation system. In this chapter an overview is given of the available R&D input data at (supra)national and regional level. With the Frascati Manual the OECD provides a methodological framework for setting up national surveys to collect these data. This methodology is used to produce standardised measurements of human and financial resources devoted to R&D by OECD member countries. EUROSTAT adapted and extended this methodology to produce for the EU countries R&D input data at regional level. To measure the performance of national and regional R&D systems, input and output data have to be combined. The methodologies for collecting input and output data have, however, been developed largely independently from each other. The resulting limitations on their use in performance indicators are discussed, and suggestions are formulated for a more integrated approach to construct input and output data.

1. INTRODUCTION

Because after the Second World War more national resources were devoted to research and development (R&D), several industrialised countries including the United States, Canada, France, and the Netherlands started to collect statistical data about these activities. Methodologically setting up R&D surveys turned out to be rather complex and differences in scope and methodology made international comparisons difficult.

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 315-338.

Within the framework of multilateral organisations operational definitions were elaborated for research and development and for the broader concept of scientific and technological activities, and statistical techniques were developed to measure private and public resources invested in these activities. For the former the OECD laid down the methodological framework in the Frascati Manual (OECD, 2002a) and for the latter UNESCO formulated the 'Recommendations concerning the International Standardisation of Statistics on Science and Technology' (UNESCO, 1978). Methodologies for generating data about R&D investment and human resources have been constantly upgraded and extended.

With the gradual emergence of the knowledge based economy over the last few decades, research and technological innovation moved centre stage. A better understanding is needed of the production, the accumulation and the dissemination of knowledge, as well as its use in innovation. Information on R&D investment, by itself, is not adequate for making an evaluation of the efficiency and impact of R&D activities. Internationally comparable data and indicators on the results of R&D in the form of new knowledge and applications and downstream new technological innovations are necessary.

With the rapid development of information technologies the analyses of information extracted from bibliographic databases of scientific publications, indicated by the generic term bibliometrics, developed into a fully fletched scientific discipline. Initially limited to information about scientific production and visibility, progressively more sophisticated multi-dimensional indicators are developed to map scientific fields and the specialisation of entities such as countries, regions, or research organisations (van Raan, 1988). Parallel with bibliometrics, the study of patent data retrieved from the most important international patent registration systems has been development to gain insight into the technological capacity and competitiveness of entities (Jaffe and Trajtenberg, 2002). Combining bibliometric and patent data leads to a better understanding of the flows of knowledge, its use, and the complexity of the innovation process (Narin et al., 1997; Verbeek et al., 2003).

Although some work has been carried out on bibliometrics (Okubo, 1997), OECD has paid more attention to develop internationally agreed standards for the use of patent data as R&D indicators (OECD – Patent Manual, 1994). To monitor the innovation process OECD produced methodological manuals on Technological Balance of Payments (OECD – TBP Manual, 1990), Innovation (OECD – Oslo Manual, 1997), and S&T personnel (OECD – Canberra Manual, 1995). Some of these Manuals were produced in close collaboration with EUROSTAT, the European Union's statistical office. Together with the Frascati Manual on R&D input data and the Patent Manual, these three manuals form the Frascati Family.

The methodologies for collecting on the one hand input data and on the other hand output data on national or regional R&D systems have been developed largely independently from each other. However, to measure the productivity of R&D systems, input and output data have to be combined (Bonaccorsi and Daraio, this Handbook). The outcome of this process depends critically on an integrated approach to generate the underlying data. In this contribution, a review is made of the available input indicators and the potentialities and pitfalls of their use in performance analyses of R&D systems.

In Section 2 an overview is given of the R&D input data published by the most important international organisations and the methodological framework for generating them. In section 3 input data published in indicator reports of a selected number of countries are presented. Section 4 describes the potentialities and pitfalls in the use of input data to compare the performance of national and regional R&D systems.

2. INTERNATIONAL STANDARDS FOR R&D INPUT DATA

In this section an overview is given of the efforts of the OECD and UNESCO to establish a methodology to collect internationally comparable input data.

Using these techniques, and often extending or fine-tuning them to their specific needs, supra-national organisations in Europe, Asia, and Latin America also collect input data to integrate them into monitoring tools to assess the performance of their scientific and technological capacity. The European Union through its statistical arm EUROSTAT has made a huge effort. In this section some of EUROSTAT's work is presented. Also the South East Asian countries through the Pacific Economic Cooperation council (PECC) and Asia Pacific Economic Cooperation Council (APEC) networks and Latin American countries through the Ibero American Network on S&F indicators (RICYT) are catching up fast.

2.1 OECD's Frascati Manual[1]

2.1.1 A brief history

To measure activities it is first and foremost necessary to agree on the study's object. As far back as the 1960s the OECD set up a working group composed of national experts to clarify the concepts and elaborate specific proposals for the standardisation of R&D statistics.

The 'Proposed Standard Practice for Surveys of Research and Development' was developed to provide guidelines for OECD member countries to collect and issue national data on R&D expenditures and on R&D personnel and to submit responses to OECD R&D surveys. It was accepted by national experts from the OECD member countries at a conference, held in Frascati, Italy, in June 1963 and became known as the 'Frascati Manual'.

After the adoption of the guidelines, the OECD launched in 1964 an International Statistical Year on R&D. Seventeen countries took part in this exercise. Following the publication of the Statistical Years findings, the OECD revised the manual taking into account suggestions of member countries and making it conform, as far as possible, to existing United Nations' international standards. OECD published the revised version of the Frascati Manual in 1970

In the second revision, adopted in 1974, the scope of the Manual was expanded to cover research in humanities and social sciences and the 'functional' classification by 'funding objectives' was elaborated. In 1980 a fourth edition was published with only minor revisions.

The fifth and sixth editions revised the guidelines to take into account developments in the S&T systems and the new insights in it. The former took into account issues such as internationalisation, the role of software, and transfer sciences. The rationale for setting up the fifth and latest revision of the Frascati Manual included the need to update various classifications and an increasing need for data about R&D in the service sector, about globalisation of R&D, and about human resources for R&D.

Applying the methodology laid down in the Frascati Manual, OECD member countries set up surveys to collect national R&D input data and submit the responses to this organisation. OECD standardises these data and

[1] The information in this subsection is based on the 2001 edition of the Frascati Manual (OECD, 2002b). From the next edition onwards, OECD decided to change this publication's title in "Research and Development Statistics".

publishes the information annually in the 'Basic Science and Technology Statistics' series.

2.1.2 Methodological framework

The Frascati Manual deals only with the measurement of research and experimental development. OECD defines 'Research and Experimental Development' (R&D) as follows:

> " 'Research and experimental development' comprise creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture, and society, and the use of this stock to devise new applications".

To operationalise this definition for the purpose of collecting data the Frascati Manual enumerates criteria for distinguishing R&D from related activities and describes in detail the activities not considered as R&D, particularly education and training, and administration and other supporting activities.

For the purpose of the surveys the OECD defines the lowest aggregation level or statistical unit as the entity for which the required statistics are compiled. It can be either an observation unit from which the observation is received or an analytical unit which statisticians create by splitting or combining observation units with the help of estimations or imputations in order to supply more detailed and/or homogeneous data then would otherwise be possible.

Ideally the statistical unit should be uniform, within sectors, for all countries. In practice this goal is never fully achieved owing to national differences in the structure of the R&D organisations. Moreover, from sector to sector and from country to country the reporting unit, from which the data are collected, may differ from the statistical unit. If the reporting unit is larger than the statistical unit difficulties may arise for distributing the data among the appropriate statistical units.

To make internationally comparable analyses of R&D input data the Frascati Manual recommends the use of a standardised *institutional classification* and a *functional classification* based on the nature of the R&D activities of the statistical units

The institutional classification is based on the principal economic activity of the statistical unit. This classification which follows as closely as possible the System of National Accounts (SNA). The 2001 edition of the Frascati

Manual uses the 1993 version of the SNA [2]. It has five main (economic) sectors: business enterprise sector; higher education sector; government sector; private non-profit sector; and the sector 'Abroad'.

The *business enterprise sector* includes firms, organisations, and institutions whose primary activity is the market production of goods or services (other than higher education) for sale to the general public at an economically significant price, and private non-profit institutions mainly serving them. For the international comparisons of R&D statistics the Frascati Manual recommends that statistical units in the business enterprise sector be classified in industrial groups and subgroups in which it has its principal activity or range of activities. This classification is based on the most recent version of the International Standard Industrial Classification (ISIC Rev 3.1, UN, 2002)[3]. Enterprises should also be classified by type into three organisational categories (private, public, and a residual category 'Other research and co-operative institutes') and by size, preferably based on the number of employees.

Different from the SNA 1993, the Frascati Manual establishes the *higher education sector* as a separate sector. This sector is composed of all universities, colleges of technology and other institutions of post-secondary education, whatever their source finance or legal status. It also includes all research institutes, experimental stations and clinics operating under the direct control of — or administered by, or associated with — higher education institutions. In all countries a large share of R&D is performed in this sector. However, this sector is not included in the SNA and the above definition is susceptible to variations in interpretation, resulting in difficulties in obtaining internationally comparable data. The core of this sector is formed by universities and colleges of technology, the classification problems are mainly in the periphery of the sector with post-secondary institutions and institutions linked to universities such as hospitals, clinics, and mission oriented research institutes managed by — or affiliated with — universities.

[2] SNA consists of a coherent, consistent, and integrated set of macro-economic accounts, balance sheets and tables based on a set of internationally agreed concepts, definitions, classifications and accounting rules. It provides a comprehensive accounting framework within which economic data can be compiled and presented in a format that is designed for purposes of economic analysis, decision taking, and policy making. It also serves as a point of reference in establishing standards for related statistics.

3 ISIC is the United Nations International Standard Industrial Classification of All Economic Activities. This classification is the international standard for the classification of productive economic activities. The main purpose is to provide a standard set of economic activities so that entities can be classified according to the activity they carry out.

The statistical units in the higher education sector are to be classified in six major fields of science and technology, proposed by UNESCO's 'Recommendations Concerning the International Standardisation of Statistics on Science and Technology' (UNESCO, 1978): Natural sciences; Engineering and technology; Medical sciences; Agricultural sciences; Social sciences; and Humanities.

Although examples are given of sub-fields for each major field, no recommendations about a classification scheme are formulated. The Frascati Manual states:

> "While the major fields of science and technology are clearly defined, the level of desegregation within each component is left to each country's discretion."

To collect the data an appropriate breakdown of the organisations into smallest homogeneous statistical units with each unit's principle activity in only one of the six fields has to be made. However, the lack of a clear delimitation or description of each major field makes this classification somewhat arbitrary and hampers the international comparability of the data at the disciplinary level.

The *government sector* is defined as:

– All departments, offices, and other bodies which furnish, but normally do not sell to the community, those common services, other than higher education, which cannot otherwise be conveniently and economically provided, as well as those that administer the state and the economic and social policy of the community (Public Enterprises are included in the business enterprise sector).
– Non-profit institutions controlled and mainly financed by government, but not administered by the higher education sector.

The Frascati Manual does not contain for the government sector a recommendation for an appropriate sub-classification of R&D activities. But it is recommended for classifing the statistical units into three categories, according to the level of government involved (central and federal government units, provincial and state government units, local and municipal units) along with a fourth category for units that cannot be distributed by level of government. This classification is especially relevant for countries with a federal state structure as regional authorities generally play an important role in R&D policy.

The *private non-profit sector* includes:

– Non-market, private, non-profit institutions serving households (i.e. the general public).

– Private individuals or households.

By convention this sector covers the residual R&D activities of the general public, which plays a very small role in the performance of R&D. The sector ‘Abroad’ consists of:

– All institutions and individuals located outside the political borders of a country, except vehicles, ships, aircraft and space satellites operated by domestic entities, and testing grounds acquired by such entities.
– All international organisations (except business enterprises), including facilities and operations within the country’s borders.

The *sector ‘Abroad’* occurs in R&D surveys only as a funding source for R&D performed by statistical units already classified in one of the four national sectors or as a destination for extramural R&D expenditures. The Frascati Manual recommends classifying the origin or destination of these funds by geographic area with as institutional sub-classification the four sectors used for domestic R&D plus as a fifth ‘international organisations’.

Using this institutional classification based on five sectors, difficulties may arise in classifying units by sector, because it is not always clear into which an institution should be classified. Institutions may, for example, straddle two sectors. Given these difficulties, when two countries classify institutes with similar functions in different sectors the surveys produce results that are not fully comparable.

In the functional classification the R&D resources of the performing units are distributed to one or more functional classes on the basis of the characteristics of the R&D activities. In this classification four sub-classifications are distinguished:

– Type of R&D;
– Product fields;
– Fields of Science and Technology;
– Socio-economic objectives.

The Frascati Manual recommends for all four national sectors of performance the breakdown by *type of R&D*, distinguishing three types: basic research; applied research; and experimental development. Basic research is defined as experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundation of phenomena and observable facts without any particular application or use in view. Applied research is also original investigation undertaken in order to acquire new knowledge, however directed primarily towards a specific aim or objective. Experimental development is defined as systematic work, drawing on existing knowledge gained for research and/ or practical experience, which is directed to producing new materials, products or devices, to installing new processes, systems or services, or to improving substantially those already produced or installed. R&D covers both formal R&D in R&D units informal or occasional R&D in other units. The Frascati Manual contains criteria for distinguishing between the three types of research.

The distribution of R&D by *product fields* is confined to the business enterprise sector. The distribution of R&D by product fields allows a more appropriate distribution of R&D resources to the relevant industries because for the purpose of detailed analysis they are more comparable internationally. Basic research is curiosity driven and cannot be assigned to product fields, but in practice basic research carried out by a firm is generally oriented towards a field of interest of this firm, given its potential commercial applications. The Frascati Manual recommends classifying the basic research carried out by firms in their field(s) of interest. The product field distribution should be based on the International Standard Industrial Classification[4] (ISIC) (UN, 1990).

For the functional distribution of R&D resources data at project level are generally collected and a detailed list of fields should be used. However OECD has not agreed on a standard classification list of *fields of science and technology* suitable for the functional distribution of R&D activities. Until such a list is developed it is recommended that the major fields used for the institutional distributed are used.

R&D performers should retrospectively report about the primary *socio-economic objectives* of their intramural R&D activities. This breakdown is most easily applied in the government and private non-profit sectors. However even this is not systematically done by all OECD member countries. In the sixth edition of the Frascati manual OECD suggests that member countries should at least make efforts to collect performer reported data in all sectors for two priority objectives: defence; control and care of the environment. The distribution list that is recommended for the classification of R&D expenditures by socio-economic objectives is based on Nomenclature for the Analysis and Comparison of Scientific Programmes and Budgets (NABS revision 1992)[5] (except for research financed from

[4] The ISIC is intended to be the standard classification of productive economic activities. Its main purpose is to provide a set of categories which can be utilised when dissecting statistics according to such activities. It has 17 sections, subdivided in two-digit divisions, three-digit groups and four digit classes.

[5] NABS is a functional classification for the analysis of public financing of research and development (R&D) on the basis of the socio-economic objectives pursued by the central governments or stated by them in drafting their budgets and programmes, as opposed to a breakdown by institutions or groups of institutions to which funds are allocated. The

general university funds, which is not appropriate for performer based analysis).

A statistical unit may have R&D expenditures within the unit (intramural) or outside it (extramural). Intramural expenditures are defined as all expenditures for R&D performed within a statistical unit or sector of the economy during a specific period, whatever the source of funding. Both current and capital expenditures should be included. Current costs are composed of labour costs (including costs of postgraduate student at PhD level on the payroll of universities or R&D units and /or receiving external funds for R&D) and other current costs (e.g. non-capital purchases of materials and supplies necessary to carry out the research). The Frascati Manual stipulates that data on R&D expenditures should be by cost factor (excluding VAT and similar sales taxes).

Capital expenditures are the annual gross expenditures on fixed assets used in R&D programs of the statistical units. They should be reported in full for the period when they took place and should not be registered as an element of depreciation. The Frascati Manual describes in detail which expenses should be considered as intramural capital expenditures. The intramural expenditures should be classified by the source of funding as reported by the performer. To correctly identify the flow of funds the transfer must be direct and both intended and used for the performance of R&D. In surveys performers are usually asked to distribute their intramural expenditures between funds of the performing unit (own funds), funds from other units in the same sector or sub-sector, and funds from other sectors and sub-sectors.

The Frascati Manual recommends a classification scheme for the funding sources based on five sectors: *Business enterprise*; *Government*; *Private non-profit*; *Higher education*; and *Abroad*. The fraction of public general university funds (GUF) used to support R&D, which represent in most countries a large fraction of publicly funded R&D, form a sub-sector of the Government sector. The Frascati Manual also recommends constructing the regional distribution of R&D intramural expenditures.

Extramural expenditures are the sums which a unit, organisation, or sector reports having paid or committed themselves to pay to another unit, organisation or sector for the performance of R&D during a specific period. This includes acquisitions of R&D performed by other units and grants given to others for performing R&D. OECD recommends for the distribution of extramural R&D the use of a classification scheme similar to that for the

---

NABS was devised as a means of describing the appropriations of central government for research activities, and not the actual execution of the work.

intramural ones. In the context of a globalisation economy and a growing international collaboration, at the country level extramural funding of R&D activities becomes increasing important.

Based on the surveys of statistical units' R&D expenditures, a country's gross domestic expenditure on R&D (GERD), the total intramural expenditures on R&D performed on the national territory during a given period can be calculated. GERD includes R&D performed within a country and funded from abroad but it excludes payments for R&D performed abroad. The Frascati Manual states that it would be useful to have separate tables for defence and civil R&D in order to monitor how trends in these areas affect the level and structure of total GERD. On the other hand the gross national expenditure on R&D (GNERD) comprises total expenditure on R&D financed by a country's institutions during a given period. It includes R&D performed abroad but financed by national institutions or residents; it excludes R&D performed within a country but financed from abroad. To identify the R&D activities of international organisations the sector 'Abroad' should have sub-categories for international organisations.

The GERD and the GERD matrix (with the funding sector as one variable and the sector of performance as the other) are the basis for international comparisons of R&D expenditures. To compare these expenditures their values must be adjusted for differences in prices level among countries (interspatial differences). To study the evolution, the same must be done for differences in prices levels over time (inter-temporal differences). Special methods have been developed for deflating and converting data about R&D expenditures expressed in national currencies at current prices to a numeraire currency. However, a full set of R&D deflators and R&D converters are not available for all OECD member countries. Therefore the Frascati Manual recommends the use of the implicit gross domestic product (GDP) deflator and purchasing power parity for GDP (PPP-GDP), which provide an approximate measure of the average real 'opportunity cost' of carrying out the R&D work.

Expenditure data measure the total cost of carrying out the R&D, including indirect support (ancillary) activities. Personnel data measure the resources going directly to R&D activities. To count R&D personnel all persons should be taken into account who are employed directly on R&D work, as well as those providing direct services such as R&D managers, administrators, and clerical and technical staff.

The Frascati Manual provides guidelines for classifying R&D personnel by occupation and by level of formal qualification. OECD recommends its member countries to use a least the classification by occupational category, because this classification allows the best international comparisons of the number of personnel employed in R&D.

For the purpose of R&D surveys the Frascati Manual classifies the R&D personnel into three categories: researchers; technical and equivalent staff; other supporting staff. Researchers are defined as professionals engaged in the conception or creation of new knowledge, products, processes, methods, and systems, and also in the management of the projects concerned. For the two other categories the Frascati Manual also provides operational definitions. These three definitions can be linked to the broad categories for the International Classification Of Occupations version 1988 (ISCO-88), produced by the International Labour Organisation (ILO).

For the university sector, where education and research are entwined, difficulties arise in separating both elements. The two most important are the position of post-graduate students and the fraction of the staff members' working time spent on research.

Owing to differences in the labour laws, social security regulations, and more generally, the socio-economic structure, the position of post-graduate students differs considerably between OECD countries. The Frascati Manual contains guidelines for including the categories of post-graduate students in R&D personnel series.

The Frascati Manual also provides a methodology for the headcount of R&D personnel. But also the fraction of the working time spent on R&D of the staff has to be estimated. For the higher education sector, where most staff are involved in teaching and research (and often in management and socio-economic valorisation of knowledge), this estimate is tedious but also of utmost importance, because in most OECD member countries a large share of R&D, especially basic research is performed at universities. The Frascati Manual recommends surveys of the use of time to be carried out every five to ten years. It gives a definition of the working time and stipulates that the calculation of the full time equivalent R&D personnel must be based on total working time.

The staff data should be broken down by sex and age, using Provisional Guidelines on Standards International Age Classifications, published by the United Nations (UN, 1982). As R&D moves to the centre of the stage in the knowledge economy, additional internationally comparable information about R&D personnel such as nationality, country of previous resident or country of study at the highest level could be very valuable. Although mentioning it, the Frascati Manual contains no specific recommendations for collecting these data.

In all OECD member countries public authorities fund a considerable fraction of R&D activities. The most appropriate and accurate way of measuring public R&D spending is to hold surveys of the units carrying out the research. Surveys make it possible to collect detailed information about the amount spent on R&D and about the fraction financed by public authorities. The latter is known as the ‘government financed gross domestic expenditure on R&D’ (government financed GERD). Collecting statistical data on R&D is however tedious, expensive and time consuming. Therefore the Frascati Manual describes a second methodology for measuring government support for R&D using information extracted from public budgets. It is based on the identification of all budget lines involving both running costs and capital spending on R&D, and measuring or estimating the fraction spent on R&D both as current cost and as capital expenditure. These budget based data, referred to as ‘government budget appropriations for R&D’ (GBOARD) are quicker and cheaper to generate than data about the government financed GERD. However, they are less accurate, and because the budgetary cycle evolves from forecasts to actual outlays they have to be readjusted.

In most countries not only the central government supports R&D activities, but also the regional and lower authorities. The Frascati Manual recommends that for the purpose of the GBOARD central or federal government should always be included and provincial or state government only if its contribution is significant. Local government funds should be excluded. However, because the OECD is mainly interested in collecting statistical data of its member states in a standardised form, no detailed reporting of R&D funding by the different authorities in federal countries is required.

The Frascati Manual also recommends determining the distribution of the GBOARD by socio-economic objective. The purpose of the R&D project or programme should be used to classify R&D outlays, using the NABS classification categories.

There are differences between sums reported as GBOARD and government financed GERD. The FRASCATI Manual discusses these differences. The GERD only covers expenditures on the national territory whilst GBOARD also includes public R&D funds spend abroad, such as payments to international organisations. Difference may also occur because, for example, the periods covered are different or the money is finally spent by the performer a year or more later than to which it was committed by the funding authority. Moreover, for the classification of data the reporting entities’ points of view may differ: GERD and GERD objectives are based on reports of the R&D performers whereas GBOARD is based on reports of the funding authorities.

2.2 UNESCO

Whilst the Frascati Manual is aimed at collecting input data on R&D, UNESCO pioneered since the beginning of the 1960s the definition of international standards for measuring science and technology.

R&D has certain characteristics distinguishing it from the larger family of scientific and technological activities, and from the economic activities it is a part of the 'Recommendation concerning the International Standardisation of Statistics on Science and Technology' adopted by UNESCO's General Council in 1978 (UNESCO, 1978), defines the broader concept of 'Science and Technology activities' (STA). They include 'Research and Development' (R&D), 'Scientific and Technical Services' (STS), 'Scientific and technical education and training' (STET). STS covers activities in museums, libraries, translating and editing S&T literature, surveying and prospecting, testing and quality control, and so on. STET refers to S&T education and training, notably in tertiary education. UNESCO and OECD have been using the same basic definitions for the coverage of the financial and human resources devoted to R&D.

UNESCO published three editions of the World Science Report (1993, 1996, 1998) and produces reports such as 'The state of science and technology in the world 1996 1997' (UNESCO, 2001). These publications give an overview of the available input and output data on S&T. UNESCO's Statistical Yearbook also contains information on R&D expenditures and staff and provides an estimate of the world's total expenditures on R&D. Especially with the report 'The state of science and technology in the world 1996 1997', the UNESCO Institute for Statistics made a considerable effort to collect and analyse available S&T data. Moreover, this institute published recently a strategic plan for improving the relevance, availability and quality of S&T statistics (UNESCO, 2003).

There is, however, a paucity of statistics about science and technology, especially in the developing countries, making the interpretation of the estimates difficult.

2.3 EUROSTAT

In the second half of last century, supra-national organisations set up by regional groupings of countries have become an intermediate level between the nation state and the multilateral organisations such as the OECD. The most striking example is the European Union.

To support its own policy making and provide those involved in S&T policy with reliable indicators and comparative analysis of S&T trends in Europe, the European Commission has published periodically since 1994 a report on science and technology indicators. Over a period of ten years these successive reports reflects the changes in research and innovation policies as well as the better understanding of the research and innovation system and the capacity to measure and analyse it.

There is a partial overlap between the data about R&D expenditures and staff published in EU reports on science and technology indicators and those on the OECD Basic Science and Technology Statistics. However, over the years EUROSTAT, the statistical arm of the European Commission, developed methodologies for collecting additional information about the EU's science and innovation system. Particularly well developed are data about higher education graduates and about human resources in science and technology. When available these EU data are compared with those of Japan and the US, the two other members of the triad. The 2003 edition of the indicator report (European Commission, 2003) contains a considerable amount of data about brain gain brain drain, and, more generally, about the migration of the highly skilled. In this analysis a distinction is made between two groups: foreign students enrolled in the higher education system of specific countries and foreign researchers and other personnel employed in S&T.

From the available data an overall picture emerges, by nationality, of the number of foreign students enrolled in higher education in the EU. However, no disciplinary breakdown of these numbers is available. Only for the UK at master and PhD-level and France at graduate level are data about the nationality of students enrolled in S&T fields available.

The Third European Report on Science and Technology Indicators presents also data on the R&D workforce in the EU derived from EUROSTAT's Community Labour Force statistics Survey (CLFS). This survey covers the whole working population with an occupational classification based on the International Standard Classification of Occupations (ISCO-88) and a classification by economic activity using NACE rev 1, those with university level education, and those in senior scientific and technical posts (scientists, engineers, or technicians, including teachers in higher education). These data provide longitudinal information about mobility between EU member states and also about non-EU nationals working in EU member states. CLFS gives information about (trends in) the number of non-native employees in S&T by country of origin. The CLFS data are broken down a regional level using the NUTS classification[6]. To be

[6] EUROSTAT's Nomenclature of Territorial Units for Statistics (NUTS) is a hierarchical classification providing a single uniform breakdown of each Member Country in territorial units for the production of regional statistics for the European Union.

used as input data in S&T studies, the CLFS data have several limitations. For example, there is no organisational classification of these data and no breakdown by scientific discipline. Based on the NUTS classification, EUROSTAT also publishes for the EU member countries the regional breakdown of other input data, such as expenditures in R&D (EUROSTAT, 2003).

To support the European Union and its member states' innovation policies, EUROSTAT also set up the Community Innovation Survey (CIS). This survey is based on the Oslo Manual developed jointly by OECD and EUROSTAT. The CIS has been carried out periodically since 1992, creating a better understanding of the innovation process, the innovation's impact on the economy, and Europe's progress in the area of innovation. Innovation surveys which collect data on innovation expenditure and innovation personnel may, in the future, complement R&D data (see: http://www.cordis.lu/innovation-smes/src/studies.htm).

3. NATIONAL AND REGIONAL INPUT INDICATORS

The UNESCO and OECD member countries jointly develop methodologies for collecting only those internationally comparable data on (components of) the S&T system, they at least, in principle, engage to provide these international organisations for further standardisation and dissemination.

With the growing importance of science and technology in the emerging knowledge based economy, more and more countries, and, in federal states, regions set up statistical instruments to monitor their S&T system, and publish periodically indicator reports. These data are not limited to those the OECD recommends to provide for its statistical overview. These reports also cover topics the national or regional authorities consider sufficiently valuable for the elaboration and evaluation of their R&D policies to justify the necessary investments in methodological work and in the data collection. Although a detailed discussion of national and sub-national S&T indicator systems is beyond the scope of this section, an example of 'Non-Frascati type' of data is briefly presented.

In the United States the National Science Board publishes since 1972 biannually the Science and Engineering Indicators. These reports contain detailed information about the science and engineering workforce and about students in these disciplines. Special attention is paid to doctoral degrees delivered by higher education institutes. These data are extracted from the Doctorate Records File (DRF), a virtually complete database of research doctorate recipients from 1920 to the present (for an overview of the methodology: www.nsf.gov/sbe/srs/ssed). Each year this database is updated with data gathered from the Survey of Earned Doctorates (SED), an annual census and survey of new recipients of research doctorates in the US. The data items collected include sex, age, country of citizenship, and field of specialisation. The information gathered on the survey questionnaire has been relatively stable over its 43 year history allowing longitudinal analysis of the PhD production in the US.

As the (sub)-national data are collected and published over a long time period, these time series can be used to make trends visible. However, as they are only collected in one or a few countries, often not using the same methodology, they lack international comparability.

4. POTENTIALITIES AND PITFALLS IN THE USE OF INPUT DATA IN QUANTITATIVE S&T STUDIES

Benchmarking the performance of research entities is one of the science and technology studies' prime objectives. A distinction must be made between studies at the level of (groups of) countries and regions and studies focussed on a single or a small number of organisations and/or their components such as divisions, faculties, departments, and research groups. For the latter the relevant input data with the required specifications are generally provided by the organisation commissioning the study. The output data are either provided by this organisation or generated by linking the input data to the bibliometric or patent databases. The amount of data is mostly small enough to make a rigorous quality control possible. For studies at the meso- and micro-level a more 'coarse grained' approach can be used by collecting input data from publicly available sources, such as web pages, databases about firms, bibliographical and patent databases. The latter approach is often necessary in benchmarking exercises in which characteristics of the organisation commissioning the study have to be compared with those of similar organisations which are not participating in the study. However the results are mostly only indicative owing to the poor quality of the input data that may be outdated or incomplete.

Bibliometric and patent studies comparing characteristics of regions, countries and supra-national groupings of countries relay nearly always for input data on the information provided by organisations such as OECD or EUROSTAT. The surveys based on the Frascati Manual are unique, in the sense that they are exclusively set up to collect R&D input data. Other R&D

input data are often extracted from existing data sources, and rearranged for the purpose of science and technology studies. Census, labour force, and migration data, for example, are collected for other purposes and require a substantial amount of methodological work before they can be used in S&T studies. Output data extracted from bibliographical and patent databases have the same limitations.

Although great progress has been made in collecting internationally comparable R&D input data and their use in S&T studies, a number of pitfalls remain. In this section some of these difficulties will be discussed.

4.1 Missing Values in Time Series or Time Series not Available for a Country or a Number of Relevant Countries

Although OECD member countries commit important resources in setting up surveys based on the Frascati Manual, for many variables the time series have missing values and/or data become available only after several years' delay or as estimated values. In the Basic Science and Technology Statistics OECD often retroactively adjusts values. Although publication databases are updated weekly, monthly, or at least yearly, it is not possible to calculate indicators using output and input data for the most recent years, or it can only be done on an ad hoc basis using values estimated with appropriate statistical techniques. For patent data the situation is somewhat more complex owing to the duration of the patent granting procedures, but at least for data about patent applications the same difficulties exist.

GERD data are generally available with two years delay, and a number of countries such as Australia, Norway, and Portugal set up R&D surveys only on a biannual basis. To make a performance analysis for publicly funded research the government budget appropriations for R&D (GBOARD) could be used as a proxy, but for several OECD countries these data have the same limitations.

The available data for the GERD by sector of performance and major field of science and technology are even more limited. A number of OECD countries such as Austria, Belgium, Denmark, and France do not provide a disciplinary breakdown, not even for the two large groups, the life, natural and engineering sciences on the one hand and the social sciences and humanities on the other. Other countries such as Canada and Germany only provide data for these two groups.

For R&D personnel the time series are even less complete. In the 2001 edition of the Basic Science and Technology Statistics in the period 1990 2001 for Austria and Switzerland, for example, these data are published for only for 2 and 3 years. For the UK no data about the number of R&D personnel have been published since 1993, and for the US only the number of researchers is available but not the number of support staff. Ten OECD member states have provided since 1990 no information about the total number of R&D staff per field of science. For research personnel working in the higher education sector more countries provide this information. However, for two important players, the UK and the US, even for the higher education sector no information about the R&D staff is published.

Table 14.1. Ratio of total GERD and total number of R&D Staff for selected countries

	90	91	92	93	94	95	96	97	98	99	00	01
Australia	66		72		76		83		83		86	
Austria				99					107			
Canada	90	91	91	92	87	87	89	93	106	110		
Czech R						58	65	70	74	70	78	74
Denmark	69	71	70	71		73						
Germany		82				87	89	90	93	96		
Hungary				32	38	34	32		36	36	42	
Iceland	48	57	61			54		57		71		
Ireland	66	63	68	72	74							
Japan		94	90	87	86	92	96	101	100	101	108	
Korea						85	106	114	107	107		
Mexico				52	64	59						
Norway		75		75		74		79		82		
Poland						24	27	28	29	34		
Portugal			70			64						
Slovak R							30	38	29	27	28	29
Spain	74	75	74	71		62		61		62		
Sweden		89				101						

For the life, natural and engineering sciences the ratio of the total GERD in $1000, expressed in constant 1995 $ prices corrected for Purchasing Power Parity and the total number of R&D staff, in FTE. Only OECD member countries with a reported value for both variables in at least one year during the period 1990 2001 are listed. The data are extracted from the 2001 edition of the Basic Science and Technology Statistics (OECD, 2002c).

Table 14.1 illustrates the limitations of the available data. Based on the 2001 edition of OECD's Basic Science and Technology Statistics, it gives by member country for the life, natural, and engineering sciences the ratio of the total GERD, expressed in constant US$ 1995 prices corrected for Purchasing Power Parity on the one hand and the number of R&D staff in FTE, on the other in the period 1990 2001. For Belgium, Finland, France, Greece, Italy, the Netherlands, New Zealand, Switzerland, Turkey, the United Kingdom, and the United States in each year no value was available for at least one of the two variables. These countries are therefore not included in Table 14.1.

4.2 Difficulties Related to Methodological Incompatibilities between Input and Output Data

Methodological incompatibilities between input and output data hamper the construction of R&D indicators. As research and development become increasingly important in the knowledge based economy, at the level of scientific disciplines more detailed information is needed about a country's or group of countries' relative research performance. As mentioned in section 2.1, the Frascati Manual only enumerates six major fields of science and technology without providing a methodological framework for classifying R&D activities in these fields. Moreover, they are too broad for a detailed analysis for which a classification of scientific disciplines at subdisciplinary level is needed. For the *Science Citation Index* (SCI; ISI – Thomson Scientific, PA, USA) the most widely used database for bibliometric studies scientific publications (indirectly through the journals in which they appear) are classified in scientific categories. These categories can be grouped into the main disciplines of the natural and life sciences and the basic disciplines of the technical sciences. Other bibliographic databases used in bibliometric studies, such as INSPEC, also contain disciplinary classification schemes. Recently work has been done to develop this classification scheme for evaluation purposes (Glänzel and Schubert, 2003).

However, because only surveys based on the Frascati Manual are set up exclusively to collect R&D input data, and in principle are very flexible, it seems most appropriate that the OECD takes the lead in elaborating a sub-disciplinary classification scheme not only for the natural and life sciences but also for the social sciences and the humanities that should be at least compatible with the one used by ISI – Thomson Scientific. Otherwise combining input data based on the Frascati Manual and publication data at (sub-) disciplinary level will remain hazardous, and the outcome at best only indicative.

As an indicator for the technological strength of countries and regions, propensity to patent is often used. To compare countries' relative positions in different technological fields, patent data have to be related to input data. For the classification of patent data the International Patent Classification system (IPC) is used. The IPC is a hierarchical classification system comprising sections, classes, sub-classes and, groups (main groups and subgroups). OECD's Frascati Manual should also contain a classification scheme or a correspondence table to make it possible at least at the level of the IPC sections, and preferably of the IPC classes, to link input and patent data.

Table 14.1 illustrates another difficulty. For the natural, life, and engineering sciences, the ratio of the standardised GERD and the number of

R&D staff in FTE differs significantly between countries, in some cases by a factor of three. These variations are also observed between total GERD and total number of R&D staff. Consequently, using in the indicators as input data GERD or total number of R&D staff may significantly alter the outcome of a performance analysis.

The statistical assumptions used to transform the raw data into standardised measurements of human and financial resources devoted to R&D may at least to some extent cause these variations. For example, the fraction of the working time of academic staff, representing in most countries an important fraction of the total R&D staff, spent on R&D has to be estimated. This fraction evolves over time and not all countries carry out at regular intervals the recommended 'time use' surveys. In Belgium, for example, such a survey has not been carried out since the middle of the 70's. Differences in indicators based on expenditures and on corresponding staff data may also be explained by characteristics of a country's S&T system or more generally its economic structure. These differences need to be better understood and should be the subject of further research.

4.3 Valuable Input Data not Collected within the Framework of International Organisations

In a globalised knowledge economy human resources in science and technology is an important policy issue not only in OECD member states but also in developing counties such as India and China. Owing to an increase in knowledge intensity of the economy combined with in some, especially Western European countries, a rapidly ageing workforce, there is apparently a growing shortage of highly skilled workers, particularly researchers in some sectors (European Commission, 2003). This shortage is supposed to become even more critical at the end of this decade. Information on the stock and flows of researchers are needed to analyse these phenomena and to understand the performance of R&D systems. Although the OECD and EUROSTAT work on methodologies for collecting more and better data on R&D staff, the material available has to be improved and based on internationally agreed standards time series have to be produced. OECD's publication 'International Mobility of the Highly Skilled' (OECD, 2002) gives an overview of the available indicators and their shortcomings

Data about PhD degrees and about the next steps in the careers of young doctorates deserve special attention, because they are an indicator for the renewal and attractiveness of a country's science system and of the performance of its research funding. Indeed information about PhD degrees constitutes simultaneously input and output data. As indicated in section 3, the US has collected since 1920 this information systematically. Given the worldwide attractiveness of the US for young researchers, many countries use these data to make qualitative arguments about the state of their own S&T system. Only a few other countries, however, collect systematically information about PhD degrees their universities deliver.

Setting up an internationally agreed methodology for collecting data about PhD degrees and the (early) career stages of its recipients along the same lines as used in the US, and collecting the data would provide an important instrument for analysing a central element in any human resources policy for science and technology. For the European Union this is particularly urgent because it is confronted with the challenge of training the researchers it needs in order to invest efficiently the additional R&D funds that come available when the recommendation of the Lisbon Summit are progressively implemented in a context of massive retirement over the next 10 to 15 years at universities and public research institutes.

5. CONCLUSIONS

Over the last half century R&D became the driving force behind economic and societal development. In parallel with the growing importance of R&D, science and technology studies evolved into a scientific discipline. The complexity of the S&T system and downstream of the innovation processes becomes progressively better understood and new insights support the development of national and supra-national S&T policies.

As governments, but also (large) corporations, increase their R&D investments to strengthen their competitive position, they need information about the effectiveness of these efforts. The prime feedback to governments' mechanism of policy making is an extensive set of R&D indicators. They help policy makers evaluate policies and to design and implement new ones.

Measuring and comparing the productivity of R&D systems is at the forefront of work on indicators. Development of such indicators critically depends on the availability of the underlying raw data. International organisations and especially OECD and EUROSTAT have made a major effort to gradually develop a methodology for collecting standardised R&D input data. As indicated in section 2.1.1, OECD has expanded considerably the scope of the Frascati Manual, including in its latest revision recommendations for collecting data on the service sector. Ongoing work at the OECD with an important impact on R&D statistics includes the development of a methodological framework for including capitalisation of R&D in national accounts. However, several member countries do not provide OECD all the statistical information recommended in the Frascati

Manual limiting the usefulness for S&T studies of the data published in the Basic Science and Technology Statistics.

As indicated in the Frascati Manual there are several methodological pitfalls that have to be tackled in order to improve the quality of the input data. A major difficulty is the lack of consensus amongst OECD member countries on (a) classification scheme(s) of scientific disciplines that, moreover, should be compatible with the most frequently used disciplinary classifications for R&D output data. Tackling this problem is a prerequisite for progress in the development of production functions and their use in the benchmarking of countries' and regions' R&D performance at disciplinary level.

Whilst on the side of output data some efforts could still be made, they are generally constrained by the rigidity of the data sources because they are not created for the purpose of S&T studies. Input data are more flexible because specific surveys are set up to collect them. This flexibility should be exploited to produce more fine grained input data. OECD and the World Intellectual Property Organisation (WIPO) recently started analysing this approach for indicators based on patent data. The same effort should be made for bibliometric data.

At the same time new data on the R&D workforce are needed in an internationally standardised format as gender, international mobility and brain drain brain gain become important policy issues. Better and coherent data on PhD degrees are a priority, because all countries recognise the importance of young doctorate holders in the renewal of the S&T system and its innovative capacity.

ACKNOWLEDGMENT

The author thanks Dr. H.F. Moed for stimulating discussions and useful comments on the manuscript.

REFERENCES

Bonaccorsi, A., Daraio, C. (2004). Econometric approaches to the analysis of productivity of R&D systems. Production functions and production frontiers. This Handbook.

European Commission (2003). Third European report on science & techmology indicators 2003. Towards a knowledge–based economy. Brussels: European Commission, Directorate-General for Research.

Eurostat (2003). Statistics on science and technology, 1980-2002 (CD-ROM).

Glänzel, W., Schibert A. (2003). A new classification scheme of fields of science and subfields designed for scientometric evaluation purposes. Scientometrics, 56, 357–367.

Jaffe, A.B., Trajtenberg, M. (2002). Patents, citations and innovations: a window on the knowledge economy. Cambridge: MIT Press.

Narin, F.J., Hamilton, F., Olivastro, D. (1997). The increasing linkage between US technology and science. Research Policy, 26, 317–330.

OECD (1990). Proposed standard method of compiling and interpreting technology balance of payment data TBP Manual 1990. The measurement of scientific and technological activities series. Paris.

OECD (1994). Using patent data as science and technology indicators Patent Manual 1994. The measurement of scientific and technological activities series, Paris.

OECD (2002). Basic science and technology statistics (CD-rom).

OECD (2002). International mobility of the highly skilled.

OECD (2002). Frascati Manual. Proposed standard practice for surveys on research and experimental development.

OECD / EUROSTAT (1997). Proposed guidelines for collecting and interpreting technological innovation data Oslo Manual. The measurement of scientific and technological activities series, Paris.

OECD / EUROSTAT (1995). The measurement of human resources devoted to science and technology — Canberra Manual. The measurement of scientific and technological activities series, Paris.

Okubo, Y. (1997). Bibliometric indicators and analysis of research systems, methods and examples. OECD, STI Working Paper 1997/1.

UNESCO (1978). Recommendations concerning the international standardisation of statistics on science and technology.

UNESCO (2003). Immediate, medium and longer term strategy in science and technology statistics. Internal review of science and technology statistics and indicators (www.uis.unesco.org).

UNESCO, (2001). The state of science and technology in the world, 1996–1997. (www.uis.unesco.org)

van Raan, A.F.J. (Ed.). (1988). Handbook of quantitative studies of science and technology. Amsterdam: North-Holland.

Verbeek, A., Debackere, K., Luwel, M. (2003). Science cited in patents: A geographic "flow" analysis of bibliographic patterns in patents. Scientometrics, 58, 241–263.

Chapter 15

METHODOLOGICAL ISSUES OF WEBOMETRIC STUDIES

Peter Ingwersen and Lennart Björneborn

Department of Information Studies, Royal School of Library and Information Science, Copenhagen, Denmark. E-mail: pi@db.dk

Abstract: The contribution defines webometrics within the framework of informetric studies, bibliometrics, and scientometrics as belonging to library and information science, and associated with cybermetrics as a generic sub-field. It outlines a consistent and detailed link typology and terminology and makes explicit the distinction between the web node levels when using the proposed terminological structures. Secondly, the contribution presents the meaning, methodology and problematic issues of the central webometric analysis types, i.e., Web engine and crawler coverage, quality and sampling issues. It discusses briefly Web Impact Factor and other link analyses. The contribution finally looks into log studies of humanWeb interaction.

1. INTRODUCTION: COLLECTION METHODOLOGY FOR WEBOMETRIC DATA

Since the mid-1990s escalating efforts have been made to study the nature of the World Wide Web, named the Web in this article, by applying modern informetric methodologies to its space of contents, link structures, and search engines. Studies of the Web have been named 'webometrics' by Almind and Ingwersen (1997) or 'cybermetrics', as in the electronic journal of that name[1]. This contribution points to research methods applied to

[1] http://www.cindoc.csic.es/cybermetrics/

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 339-369.

selected areas of webometric investigations. These areas are search engine and Web crawler coverage, quality and sampling issues; link analyses, including Web impact analysis, and log studies of Web interaction behavior. The contribution is not an exhaustive review, but rather a view of the specialty.

Webometrics displays several similarities to informetric and scientometric studies as well as the application of common bibliometric and informetric methods. For instance, simplistic counts and content analysis of web pages can indeed be seen as analogous to traditional publication analysis; counts and analyses of outgoing links from web pages, here named outlinks, and of links pointing to web pages, called inlinks, can be seen as somehow similar to citation analyses. Outlinks and inlinks are then regarded like references and citations, respectively, in scientific articles. However, since the Web consists of contributions from anyone who wishes to contribute, its quality of information and knowledge is opaque owing to the lack of peer reviewing. Hence the Web most frequently demonstrates web pages of non-scientific nature or contents. An additional difference from traditional scientific databases and archives is the dynamics of the Web, i.e., web pages and entire sites may frequently alter contents, link structure, or completely disappear. Further, the links are *not necessarily normative*, such as credit granting or recognition providing devices, but rather *functional*, say, navigational in nature. There exists no convention of linking as in the scientific world. Further, *time* plays a different role on the Web, e.g., links can be deleted, and simultaneous reciprocal linking is a rare case in the conventional citation world and not possible in the paper based scientific communication. The analogy between links and references or citations is hence of the superficial kind and should definitively not be taken too far. On the other hand, the same analogy may indeed provide interesting hypotheses about the characteristics of links and their meaning. Also, the coverage of search engines of the total Web can in principle be investigated in the same way as the coverage of domain and citation databases in the total document landscape and possible overlaps between engines can be detected. Patterns of Web search behavior can be investigated as in traditional information seeking studies. Issue tracking and mining on the Web is feasible and knowledge discovery can be carried out, similarly to common data or text mining in administrative or textual (bibliographic) databases.

Because the Web is a highly complex distribution of all types of information carriers produced and searched by all kinds of people it is central to investigate as a social phenomenon; and informetrics indeed offers some methodologies to start from. However, one must be aware that *data collection* on the Web depends on the retrieval features of the various search engines and web crawlers or robots. Although their consistency has improved from the mid-1990s, as demonstrated by Rousseau (1997; 1999), the various Web engines do not index the entire Web, their overlaps are not substantial (Lawrence and Giles, 1998), and their retrieval features are often too simplistic for extensive webometric analyses online. Sampling becomes thus an important issue, but is difficult to perform in a controlled manner.

The contribution first defines webometrics within the framework of informetric studies, as belonging to library and information science, and associated with cybermetrics as a generic sub-field. It outlines a link typology and terminology and makes a distinction between the web node levels when carrying out link analyses. Secondly, methods and methodological problems for Web engine coverage and quality studies, including Web crawling and sampling are discussed. This is followed by link analysis issues, including Web Impact Factor (WIF) analysis, and studies of Web interaction. The contribution owes substantially to recent works by Björneborn and Ingwersen (2001, 2004); Thelwall, Vaughan and Björneborn (2005); and Björneborn (2004).

2. THE FRAMEWORK OF WEBOMETRICS AND LINK TERMINOLOGY

Webometrics and cybermetrics are currently the two most widely adopted terms in library and information science (LIS) for this emerging research field. They are generically related, see Figure 15.1, but often used as synonyms. As originally in Almind and Ingwersen (1997), the present contribution distinguishes between studies of the Web and studies of *all* Internet applications. In this novel framework by Björneborn (2004) and Björneborn and Ingwersen (2004), *webometrics* is defined as:

"The study of the quantitative aspects of the construction and use of information resources, structures and technologies on the Web drawing on bibliometric and informetric approaches."

Hence, this definition covers quantitative aspects of both the construction and the usage sides of the Web taking on four main areas of current webometric research: (1) web page content analysis; (2) web link structure analysis; (3) web usage analysis (including log files of users' searching and browsing behavior); (4) web technology analysis (including search engine performance). All four central research areas include longitudinal studies of changes on the Web of, for example, search engine coverage, page contents, link structures, and usage patterns. In this webometric context the concept of

[2] A recent textbook emphasizing this point view is Hand, Mannila & Smyth (2001).

web archaeology (Björneborn and Ingwersen, 2001) is regarded as important for recovering historical web developments, for instance, by means of the Internet Archive (www.archive.org).

The above definition places webometrics as a LIS specific term in line with bibliometrics and informetrics, such as done by Cronin (2001). The terms 'drawing on' in the definition denote a heritage without limiting further methodological developments of web–specific approaches.

In the present framework, cf., Figure 15.1, *cybermetrics* is proposed as a generic term for:

"The study of the quantitative aspects of the construction and use of information resources, structures, and technologies on the *whole* Internet drawing on bibliometric and informetric approaches" (Björneborn, 2004; Björneborn and Ingwersen, 2004).

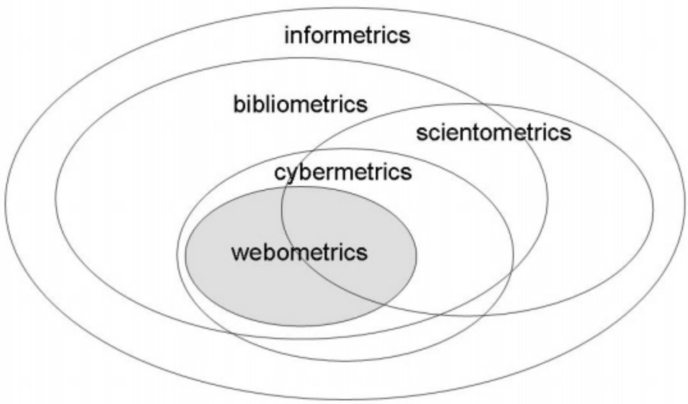


Figure 15.1. The relationships between the LIS fields of infor-/biblio-/sciento-/cyber-/webo-/metrics. Sizes of the overlapping ellipses are made for sake of clarity only (Björneborn,2004).

Cybermetrics thus encompasses statistical studies of discussion groups, mailing lists, and other computer mediated communication on the Internet (e.g., Bar-Ilan, 1997; Herring, 2002), including the Web. This definition of cybermetrics also covers quantitative measures of the Internet backbone technology, topology and traffic (cf., Molyneux and Williams, 1999). The breadth of coverage of cybermetrics and webometrics implies overlaps with approaches based on propagating computer science in Web analyses of various kinds. Björneborn (2004) and Thelwall, Vaughan and Björneborn (2005) provide comprehensive details on such analysis facets.

There are different conceptions of informetrics, bibliometrics and scientometrics. The diagram in Figure 15.1 shows the field of informetrics incorporating the overlapping fields of bibliometrics and scientometrics following the widely adopted definitions by, e.g., Brookes (1990), Egghe and Rousseau (1990) and Tague-Sutcliffe (1992). Tague-Sutcliffe states that *informetrics* is "the study of the quantitative aspects of information in any form, not just records or bibliographies, and in any social group, not just scientists" (1992). *Bibliometrics* is defined as "the study of the quantitative aspects of the production, dissemination, and use of recorded information" and *scientometrics* as "the study of the quantitative aspects of science as a discipline or economic activity" (ibid.). In Figure 15.1 politico–economical aspects of scientometrics are covered by the part of the scientometric ellipse lying outside the bibliometric one. Further, the figure shows the field of webometrics entirely covered by bibliometrics. This is because web documents, whether text or multimedia, are *recorded* information stored on web servers, cf., Tague-Sutcliffe's definition of bibliometrics (1992). The recording may be temporary only, just as not all paper documents are properly archived. Webometrics is partially covered by scientometrics because many scholar activities today are web–based whilst other such activities are even beyond bibliometrics, i.e., non-recorded, such as person–to–person conversation. Webometric studies clearly also circumscribe other social domains than the scientific one.

In the diagram the field of cybermetrics exceeds the boundaries of bibliometrics, because some activities in cyberspace commonly are not recorded, but communicated synchronously, as in chat rooms. Cybermetric studies of such activities still fit in the generic field of informetrics as the study of the quantitative aspects of information 'in any form' and 'in any social group', as stated above by Tague-Sutcliffe (1992). The inclusion of webometrics opens up the fields of bibliometrics, scientometrics, and informetrics, as webometrics inevitably will contribute with further methodological developments of web–specific approaches to the development of these embracing fields.

2.1 Link Terminology and Analysis Levels

Emerging fields like webometrics inevitably produce a variety in the terminology used. For instance, a link received by a web node has been named, e.g., incoming link, inbound link, inward link, back link, and 'sitation'; the latter term coined by Rousseau (1997) amongst others, with clear connotations to bibliometric citation analysis. The term 'external link' is an example of a more problematic terminology owing to its two opposite meanings: 1) as a link pointing out of a web site or 2) a link pointing into a site.

Figure 15.2 presents an attempt to create a consistent basic webometric terminology for link relations between web nodes (Björneborn, 2004; Björneborn and Ingwersen, 2004; Thelwall, Vaughan and Björneborn, 2005). The figure implies that the Web can be viewed as a directed graph, using a graph–theoretic term (e.g., Kleinberg et al., 1999). In such a web graph web nodes are connected by directed links. The proposed basic webometric terminology in the legend of Figure 15.2 originates, hence, from graph theory but adheres also to social network analysis and bibliometrics (Otte and Rousseau, 2002; Park and Thelwall, 2003).

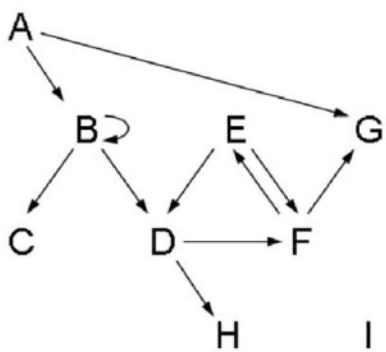


Figure 15.2. Basic webometric link terminology (Björneborn, 2004). The letters may represent different web node levels, for example, web pages, web directories, web sites, or top level domains of countries or generic sectors

- B has an *inlink* from A; B is *inlinked*; A is *inlinking*; A is an *in-neighbor* of B.
- B has an *outlink* to C; B is *outlinking*; C is an *out-neighbor* of B.
- B has a *selflink*; B is *selflinking.*
- A has no inlinks; A is *non-linked.*
- C has no outlinks; C is *non-linking.*
- I has neither in- nor outlinks; I is *isolated.*
- E and F have *reciprocal links*; E and F are *reciprocally* linked.
- D, E and F all have in- or outlinks connecting each other; they are *triadically interlinked.*
- A has a *transversal* outlink to G: functioning as a shortcut.
- H is *reachable* from A by a directed *link path.*
- C and D are *co-linked* by B; C and D have *co-inlinks.*
- B and E are *co-linking* to D; B and E have *co-outlinks.*

– Co-inlinks and co-outlinks are both cases of *co-links.*

The central terms, *outlink* and *inlink*, are commonly used in computer science–based Web studies (e.g., Pirolli et al., 1996; Chen et al., 1998; Broder et al., 2000). The term *outlink* signifies that a directed link and its two adjacent nodes are viewed from the source node providing the link, analogously to the use of the term *reference* in bibliometrics. A corresponding analogy thus exists between the terms *inlink* and *citation* with the target node as the spectator's perspective. The important conception of 'external node inlinks' in Web Impact Factor (WIF) analyses hence signifies those inlinks alone that derive from sources outside the node, i.e., excluding node selflinks. The two *co-linked* web nodes C and D in Figure 15.2 with co-inlinks from the same source node are analogous to the bibliometric concept of co-citation (Small, 1973). Correspondingly, the two *co-linking* nodes B and E having co-outlinks to the same target node are analogous to a *bibliographic coupling* (Kessler, 1963). The term 'co-links' is proposed as a generic term covering both concepts of co-inlinks and co-outlinks. The underlying assumption for the use of both the bibliometric and webometric concepts is that two documents (or two authors/link creators) are more similar, i.e., more semantically related, the higher the frequency of shared 'outlinks' (references) or shared 'inlinks' (citations).

A further discussion of this terminology can be found in Björneborn (2004) and Björneborn and Ingwersen (2004).

2.2 Levels of Link Analysis

The Web can be studied at different granularities employing what might be called micro, meso, and macro level perspectives (Björneborn, 2004; Björneborn and Ingwersen, 2004). The level depends on which of the four basic web node levels is investigated: web pages, web directories, web sites and country or generic top level domains, TLDs. Sublevels within each of the four basic node levels can exist. For example, a sub-TLD is often a central unit of analysis since many countries have assigned a level to educational, commercial, governmental and other sectors of society, for instance, *ac.uk, .co.uk, .ac.jp, .edu.au.*

Micro level webometric analyses are studies of the construction and use of web pages, web directories, and small sub-sub-sites, etc., for example, constituting individual web territories. Meso level webometrics is correspondingly concerned with quantitative aspects of larger sub-sites and sites. Macro level webometrics comprises studies of clusters of many sites, or focuses on sub-TLDs or TLDs. Several webometric studies, including classic ones by Larson (1996) and Almind and Ingwersen (1997), have used meso level approaches concerned with site–to–site interconnectivity as well as macro level TLD to TLD analyses; primarily applying *page level* link counts to all analyses. However, in order to extract useful information, links may also be *aggregated* on different node levels as in the recently developed *Alternative Document Model* (ADM) (Thelwall, 2002; Thelwall and Harries, 2003). In contrast to the classic studies, the ADM may operate, say, at a sub-site level as analysis unit (representing, for instance, university departments), and with link counts also at sub-site levels, i.e., aggregating the page level link counts. It should be noted that a site level link *always* connects a source site with a target site. Correspondingly, a page level link always connects a source page with a target page. However, a target URL for a web page may misleadingly look like an URL for a web site, since it is common web practice to stem the target URL of top entry pages of a web site. For instance, instead of writing the full URL 'www.db.dk/default.htm' in a target link pointing to the top entry page of the Royal School of LIS, it is more expedient to stem the URL to 'www.db.dk' since web servers automatically look for default pages for stemmed URLs. However, this stemmed URL still denotes a web page and not a web site.

An adequate terminology for aggregated link relations should capture both the link level under investigation and the reach of each link. Such a terminology must reflect at least three elements: (1) the investigated link level; (2) the highest level web node border crossed by the link; and (3) the spectator's perspective, i.e., do we talk about inlink or outlink analyses. As a consequence, selflinks are used for a wider range of purposes on the Web than self–citations in the scientific literature. Page selflinks point from one section to another within the same page. Site selflinks (also known as internal links) are typically navigational pointers from one page to another within the same web site. Within the same TLD individual links connecting sub-TLDs as in- and outlinks may thus be aggregated into TLD selflinks. The unit of analysis is hence a central issue in webometrics.

3. WEB ENGINE COVERAGE, CRAWLER LIMITATIONS, AND SAMPLING ISSUES

Fundamentally, data collection made by commercial search engines (or indeed also by personal crawlers or robots) for webometric analyses takes four forms defined by two dimensions. The first dimension focuses on the Web *data types* used as the starting point for retrieval: searching for known Web locations by means of URLs (such as 'known item' searching in information retrieval); or searching for some topic(s), or other content data that define the subject area for which the Web space is to be investigated.

The second dimension deals with the *retrieval strategy* applied by the search engine (or crawler): 'content crawling' to retrieve all unique Web documents; and 'link crawling', i.e., to follow the link associations between web pages. This strategy will also retrieve duplicates (cf., Thelwall, Vaughan and Björneborn, 2005). The objective of the Web analysis determines the mode of data collection. Common webometric analysis objectives are studies of:

– *Selected Web spaces*, defined by, e.g., specific institutions, subject areas, web document/page genre, or (sub-) TLDs and/or geo- locations, or specific personal names or single web sites. Analysis units can be Web pages, (sub-)sites, sub-TLDs and/or link structures or types. The studies are often descriptive analyses of Web characteristics. Data collection is either generated by sets of URLs associated with institutions or other known entities (Thelwall, Vaughan and Björneborn, 2005) or made via searching on defined search keys, like terms and keywords, personal names or other metadata (Bar-Ilan, 2001, 2002; Jepsen et al., 2004). In both cases sampling may be mandatory owing to the size of the space investigated;
– *Web indicators*, calculated by a number of Web parameters, e.g., number of inlinks to single or sets of web pages, (sub-) sites, (sub-) TLDs divided by number of web pages receiving them (i.e., a kind of WIF); outlinks, selflinks and other types of linking are potential parameters to be considered, as are genre, subject matter, locations, scientific citations received or references made, terms applied, institutions mentioned, numbers of faculty staff, etc. data collection is carried out as for selected Web spaces above, combined with means to obtain well defined numeric data on links and other parameters, like the use of specific search engine commands or Web crawlers;
– *Human actor – Web interaction*, that is, studies of generation of Web contents, architecture and structure or link motivation, searching the Web by various populations, in diversities of subject matter and domains and for a multitude of purposes. Data collection is made from Web engine logs and/or *in situ* observations of interactive activities in real-time – also over longer periods. This kind of studies is closely associated with interactive information seeking and retrieval studies in context (Ingwersen and Järvelin, forthcoming).

For all three kinds of Web studies the investigations may take place as longitudinal studies.

Obviously, if commercial search engines are used the coverage and the qualities of the retrieved and downloadable material at search time are crucial parameters for the resulting analysis. Hence coverage studies are central to webometric research.

3.1 Commercial Web Engine Coverage

Lawrence and Giles (1998) provided a substantial contribution with respect to the commercial search engine coverage of the Web space by introducing the concept of the publicly 'indexable Web'. The concept signifies the portion of the Web, which can be indexed by engines, excluding documents from commercial Web databases, such as, Dialog and the closed archives of publishers. That part of the information space is commonly called the 'hidden Web'. Lawrence and Giles (1999) also demonstrated that the coverage of any one engine is significantly limited by indexing only up to 17 % of the indexable Web. Central reasons behind this phenomenon are, for instance, the depth (exhaustiveness) of indexing at the local servers visited by the engine crawlers, which depends on the site structure and organization, and the link construction. Some search engines may also have indexing strategies that depends on 'pay for inclusion'. Lawrence and Giles (1999) applied randomly sampled Internet Protocol (IP) addresses. In that way it was possible to obtain a random selection of all web sites. However, this is not advisable owing to the introduction of the virtual server capability that allows one IP address to host many domain names and one 'chief' name only.

Also Clarke and Willett (1997) addressed the evaluation methodology of Web engines. They compared AltaVista, Excite, and Lycos. In addition, that paper provides a critical assessment of earlier research and produces a realistic methodology, including relative recall measures taken from IR research. It was found that AltaVista performed significantly better than Lycos and Excite. Oppenheim et al. (2000) produced a detailed review of the evaluation of Web search engines, including a discussion of test methodologies.

Whilst many coverage and evaluation studies looked into the relevance and number of web pages (recall) at a given point in time, other critical analyses covered link page retrieval by the engines (Snyder and Rosenbaum, 1999) or carried out Web structure investigations based on *time series*. As did Ingwersen (1998), Snyder and Rosenbaum observed large variations and inconsistency, in particular concerning the AltaVista engine's link page recovery at that time. Rousseau also observed the irregularity of that engine in two longitudinal studies (1999; 2001). In (1999) he compared AltaVista with NorthernLight on a daily basis over 21 weeks during 1999. This study used the same three common single words as test queries during the evaluation period. In line with the Web growth NorthernLight, as expected, showed a steady increase of hits. However, AltaVista demonstrated large variations over time until the particular date (October, 25, 1999) when it became re-launched in a renewed and quite stable form. At that date the number of retrieved web pages increased dramatically — with this *nova*–like effect depending on the query (Rousseau, 1999) — later to drop slightly supposedly owing to the deletion of dead link pages. Rousseau used the same techniques, including (median) filtering (2001), to tracki an event on the web (the introduction of the euro). Even though that and other engines nowadays seem much more reliable (Thelwall, 2001b; Vaughan and Thelwall, 2003), and give good coverage of academic web sites (Thelwall, 2001a), their harvesting and updating algorithms, which are commercial secrets, are rarely performed at search time – but at intervals. The algorithms and commands are subject to change without notice and their advanced features are not always fully documented. Further, from a webometric research point of view, authors sending their web pages to be included in the engine's index distort Web engine coverage. In the future, pay for inclusion for commercial or simple visibility reasons may hence also bias analyses. As for the ISI citation databases the commercial Web engines display national biases in site coverage. Vaughan and Thelwall (2004) demonstrated recently that three major search engines much better covered U.S. sites than sites from China, Taiwan, and Singapore.

3.2 Commercial Web Search Engine Download Capacity

A typical method to apply to assess the coverage of Web search engines is to enter some terms, concepts, or indeed entire query profiles, as long as they are well defined, and compare to a substantial number of known Web sites that should be retrieved by the engine(s). As mentioned above Rousseau did use common words as a starting point (1999) and Bar-Ilan, for instance, used scientific domain concepts like 'informetrics' and associated terms to investigate longitudinal coverage, links and Web contents on that subject matter (1999, 2000, 2002). Similarly, Jepsen et al. (2004) applied three plant biological terms (including synonyms and spelling variations) as controlled search keys: Plant hormones; Photosynthesis; and Herbicide resistance. As with Allen et al. (1999) below the methodological idea was to observe what happens on the Web with strictly scientific topics (Plant hormones) compared more publicly and politically discussed issues (Herbicide resistance) or commonly known concepts like Photosynthesis. The goals of the study were several, amongst which three are of interest here: 1) defining the depth and overlaps of the coverage in popular Web engines (Google; AllTheweb; AltaVista); 2) their accessibility level ready for download; 3) observing the scientific quality of the accessible material. For the latter goal see Section 4.1 below.

The individual engines retrieved different proportions of the Web in identical searches. For example, the number of hits by each engine for the key query term ‘Photosynthesis’ was 238,000 (Google), 119,300 (AllTheWeb), and 79,400 (AltaVista). Although the average overlap of the accessible URLs was quite substantial, 21%, the variation was also very high, from 13% to 58%. Search engine overlaps might prove to be well suited as a quality indicator, but evidently, a union of engine results may improve drastically the recall, i.e., the amount of Web materials conceivably to be investigated further.

Furthermore, the level of accessibility varied from engine to engine: only AllTheWeb provided access to several thousands of the indexed and retrieved Web publications (4,100). Google’s cut-off was close to 1000 URLs and AltaVista only allowed access, presentation, and download of 200 URLs. Bar-Ilan emphasised (2001, p. 22) how this might also be a problem to the informetrician who is interested in the whole set of results for a given query, whereas it might be less problematic to the average user who only needs a few ‘most relevant’ URLs. Unfortunately, owing to the *skewness* found in accessible URLs between engines caused by the different page ranking algorithms applied by the engines, the data material collected may not easily legitimise a correlation analysis between search engine overlaps and quality assessed by experts. The overlaps are defined by the accessibility *and* the various ranking algorithms applied by the engines. Extended potential overlaps may hence exist between the engines outside the ranked list of URLs that can be downloaded and analysed. This facet of data collection of the Web poses problems for sampling, Section 3.4.

Notwithstanding, the reason why a webometric focus commonly is put on AltaVista is that the engine has a rather large Web coverage and hitherto has provided advanced search features fitting informetric studies of the Web. For example, AltaVista allows long and complex search strings consisting of Boolean operators combined with truncation options and specific search codes for various HTML elements. Time series paired with testing for the retrieval of the query search keys and known item (site) searching (applying AltaVista’s `host:` command), conceivably also comparing to other engines’ search results, seem thus very useful as tools when monitoring Web engine performance.

3.3 Web Crawler Issues

All of the considerations demonstrated above are incentives for researchers to develop data collection techniques that do not rely upon

AltaVista or any other search engine. This implies developing dedicated 'personal' Web crawlers or robots, i.e., software that automatically and iteratively downloads web pages and/or may mine and store their links and content (Thelwall, 2001a). The issue here is whether the software reaches all potential web pages, inlinks and outlinks and their associated remote web pages, for a given site or sites, or entire domains or genres (Björneborn, 2004) in defined locations.

Commonly a crawler does not extract 100 percent of the intended data. Fundamentally, the same problems concerning data collection and sampling methods for commercial Web engines also concern 'personal' crawlers – and *vice versa*. Problematic issues associated to data collections, leading to omissions of or invalid data are, for instance, (Thelwall, Vaughan and Björneborn, 2005):

– *Starting point comprehensiveness*, i.e., URL(s), contents search keys, metadata;
– *Crawler strategy*, i.e., content or link crawling, including 'random walks' used in graph theory (cf. Björneborn and Ingwersen, 2001);
– *Omission of web pages* because the crawler does not comply with their format, pages are protected by security measures that do not allow mining or crawling or by passwords, or servers are momentarily shot down;
– *Omission of isolated web pages*, see Figure 15.2, owing to lack of inlinks;
– *Non-integration of personal home pages* in institutional link structures within a site;
– *Crawl depth limitations* at web sites;
– *Page number limitation* per site visited;
– *Different domain names used* for the same entities under study, e.g., (inter)national corporations under .com, .net, .dk.

If an engine or a 'personal' crawler has omitted portions of a Web space, entire clusters of sites, each with links to and from the local Web space sought for as well as to other sites, do not become analysed. The result is that the original space to be investigated evidently becomes *deformed* by that omission and the resulting findings distorted. However, since the analyst may not know this phenomenon to have happened, it is commonly not taken into account.

What are not really problematic for 'personal' Web crawlers are Web-economic aspects, such as, update frequency or ranking of retrieved material. Also, after the collected material is downloaded a multitude of comparative analyses may take place, for instance, by means of combinations of Boolean sets. This is not possible in commercial engines.

However, commercial Web engines may make positive use of pages and URLs from previously indexed Web sites, including author submitted sites for visibility reasons mentioned above. An additional way in to data collection may be centered on local evaluation exercises. The Web servers under assessment may permit a total crawl of their directories, as done in academic bibliometric research evaluations of universities or defined sectors.

It seems evident that future Web analyses applying crawlers ought to explain the characteristics of the software, its well tested limitations and consequently the problematic issues encountered in the harvest of data for analysis.

3.4 Sampling Issues

The limitations of the search engines and web crawlers as well as the vastness of the web make comprehensive analyses of given Web spaces quite difficult. Sampling of links and web pages is hence necessary, either as randomised samples or in stratified systematic ways. The Thelwall (2001a) collection of downloaded UK inter-university Web links generated by means of a crawler and all known UK university URLs is an example of a very large data set from which random sampling could be made in order to make a variety of analyses. A way of obtaining web pages is to make use of the links between them. For instance, so called 'random walks' can be performed by means of a crawler starting from definite points on the Web (Henzinger et al., 2000; Björneborn and Ingwersen, 2001; Thelwall, Vaughan and Björneborn, 2005) and harvesting pages algorithmically during the walk to be part of the sample and downloaded. For instance, Hou and Zhang (2003) did experiments on two kinds of retrieval algorithms, starting from a given URL and based on either co-inlink analysis of associated web pages, or by application of linear algebra theories to find deeper relationships amongst web pages. Such modes of collection are link–dependent and is computing intensive.

Another way of generating samples of web pages via web sites is to use a commercial search engine and well-defined URLs from which a random sample has been drawn covering the space under investigation. Searching on a number of engines is required in order to obtain a list of relevant URLs as comprehensive as possible by means of comparing and pooling the results into a set of (home pages from) relevant sites. If the analysis unit is sub-sites or sub-sub-sites the sample should mirror their proportions, i.e., also take the stratification into account. If this pooled retrieval is performed by means of search keys, the different engines' ranking algorithms, as in the case below, have predisposed the accessible data. However, since several engines are applied the distortion may be decreased or neutralised.

In the Jepsen et al. study (2004) the search terms associated with the three search profiles on Plant biology were deliberately searched separately and comprehensively. This was owing to the variation in cut-offs of accessible URLs displayed by each search engine, shown above section 3.2, so that enough material was available for download, sampling and further analysis in a local software program. Since the study intended also to include other parameters of the web page contents in the analyses associated with the search keys, the accessible URLs were extracted, overlaps detected, duplicates removed and isolated and search engine distributions were investigated. This provided a pooled set of URLs for each profile that was stratified according to the distributions over engines and other parameters. From that set randomised and stratified samples were drawn; for instance, 200 web pages for each profile.

Ideally, the basic search results (number of hits) from the engines ought to have provided a qualified starting point for the creation of a stratified sample. Owing to the cut-off variation of accessibility between the engines, however, any stratification must involve the *accessible* portions of the web pages that are available. Since the real population is difficult to estimate the significance of the samples is uncertain and the results probably only valid as indications, not for generalization purposes. Rusmevichientong et al. (2001) and Thelwall, Vaughan and Björneborn (2005) have made a comprehensive review and discussion of Web sampling methods. Further Web characteristics underlying the data collection and hence influencing the analysis results are discussed below.

4. WEB PAGE ANALYSES PERTAINING TO DATA COLLECTION AND ANALYSIS

Fundamentally, two kinds of web page properties are at play influencing data isolation and processing. The first kind deals with the quality of the web pages actually retrieved and accessed. This involves the trustworthiness of the contents and the page ranking algorithms of the Web search engines. The second kind signifies properties of web pages, such as genre, number of outlinks and life span of pages and links.

4.1 Web Page Quality Analyses

Obviously the breakthrough for everybody to express themselves, practically without control from authorities, to become visible world wide, also by linking to which pages one wants to link to, to assume credibility by being 'there', and to obtain access to data, information, values and knowledge in many shapes and degrees of truth, has generated a reality of freedom of information – also in regions and countries otherwise poor of infrastructure. Although some social Web etiquettes are developing, they might not be followed. The other side of the coin is that the Web increasingly becomes a *web of uncertainty* to its users; the borderline between opaqueness, shading truth, misinformation, beliefs, opinions, visions or speculation *and* reliability, validity, quality, relevance or truth becomes increasingly thinner. Picking information becomes a matter of trust. Hence Web archaeology will in future go hand in hand with webometric analyses and methods.

Evidently the individual Web engine's ranking algorithms determine *which* web pages are the highest ranked (prioritised) and thus are accessible. Lower ranked publications cannot be accessed. In the case of Google's page–ranking algorithm (Brin and Page, 1998), for instance, the 1000 top ranked and accessible pages might disproportionately belong to particular Web genres or topics that are characterised by many inlinks from authoritative Web sources, like commercial web pages and sites. The ranking algorithms are thus central to webometric analyses, because it becomes then an issue of whether the accessible publications actually are superior in a qualitative sense, e.g., are scholar or at least not opinionated or not completely untrustworthy. In particular, Google's ranking principle, building on number of inlinks to web entities and from which kind of web pages they derive. Kleinberg's (1999) recursively defined conceptions of 'hubs' (web pages with many outlinks to authorities) and authoritative web entities (with many inlinks from 'hubs') are consequently of central interest. Does a high number of inlinks always signify (cognitive) authority or rather many other hidden characteristics, e.g., that the entity with a certain probability is a commercial site? The underlying analogy with scientific citations is conceivably at play here. As noted by Otte and Rousseau (2002), the Kleinberg approach of hubs and authorities is related to the influence weight citation measure proposed by Pinski and Narin (1976) and mimics the idea of 'highly cited documents' (authorities) and reviews (hubs) in scholarly literatures.

Quality watch and assessments are currently in high demand. In particular, the health and medical domains are important areas to investigate for such issues. For instance, Courtois and Berry (1999) observed the quality among the top-20 or top-100 ranked documents retrieved by major engines. Relevant pages were such that formally contained *all* query words. This mode of 'algorithmic (or logical) relevance = quality' is similar to the most simplistic system–driven performance measures in information retrieval research (Cosijn and Ingwersen, 2000). No expert assessments were used.

Aside from the findings the paper discusses the more or less publicly available knowledge about the different indexing/retrieval features used by any one engine.

Cui (1999) made use of citation analysis methods on the Web to detect the overlapping and high frequency inlinked (cited) sites on medical information and Allen et al. (1999) looked into the reliability (and pertinence) of bio–related web pages. The former paper refers to several other studies of health and art issues treated on the Web for which Web citation analyses have been applied as a rudimentary quality indicator. The Bradford distribution of the thousands of outlinks from the 25 top medical US Schools was used as strong ties by Cui to determine the central sites concerned with specific health topics. Allen et al.'s (1999) contribution is a survey assessed by experts of the reliability of scientific web sites. As was the case for Rousseau's longitudinal study (1999) and the Jepsen et al. study (2004) mentioned above, the survey is based on the retrieval of sites according to three exemplary queries to the NorthernLight engine on 1) 'evolution', 2) 'genetically modified organisms', and 3) 'endangered species'. For each query two experts examined the top 500 web sites sequentially until each had independently reviewed approximately 60 sites containing information pertinent to the *topic*. This assessment mode is close to the methodology used in the current worldwide TREC IR evaluation experiments, applying topicality relevance measures. From 12 to 46% of the examined top pages were deemed pertinent, dependent on the topic. The 60 pertinent sites per query were scored as 'inaccurate' if they contained factually incorrect information (av. 22%), 'misleading' if they misinterpreted science or blatantly omitted facts supporting an opposing position (av. 28%), and 'un-referenced' if they presented information without any peer reviewed references (av. 71%). The latter score is purely objective. The overall agreement values for the referees' scores for the categories of 'inaccurate' and 'misleading' were 87.8% for the 'evolution' sites, 82.8% for the 'genetically modified organism' sites, and 73.6% for the 'endangered species' web sites. Un-referenced sites accounted for more that 48% for each query (1999, p. 722).

These results lead Jepsen et al. (2004) to look for filtering mechanisms in order to be able to distinguish between academic Web material and other kinds of Web information, as initially outlined above. A sample of 200 web pages from each of the three plant biological topics was drawn and assessed by one human expert into five categories plus a class of unavailable pages (11 to 22%). The first category 'scientific' was assigned to content that was deemed to be of scientific quality, for instance, preprints, conference reports, abstracts, scientific articles; 5 to 6% of the pages belonged to that category. The second category, 'scientifically related', was assigned to materials of potential relevance for a scientific query, such as directories, CVs, institutional reports. This kind of information appeared to be abundant on the Web (17 to 25% in the analysis) and may interfere with a search for specific information on a given scientific subject. The third category of 'teaching' contained content related to teaching, e.g., textbooks, fact pages, tutorials, student papers, or course descriptions. The category accounted for 11 to 20% over the three searched topics, but was more present on web pages in Scandinavian languages (16 to 37%). The student papers did sometimes blur the distinction from formal and presumably accurate scientific papers. The fourth category 'low–grade' was reserved to content that failed to meet the criteria of the three previous groups, but still was on the topic. The category typically contained content of either commercial interest or deemed inaccurate or misleading. This proportion was definitively dependent on the nature of the topic, since 'herbicide resistance' accounted for 45% belonging to 'low grade' whilst the other two topics showed values of 15 to 27%. Content not meeting the criteria for the mentioned classes, and hence not pertinent to the topic, was assigned to the fifth category labelled 'noise' (3 to 17%).

The results of such quality assessments show that since the retrieval engines' ranking algorithms play a central role, as discussed previously, top-down sampling may produce distorted or disproportioned results that, albeit, may demonstrate something about the quality and nature of the *publicly visible Web* (Allen et al., 1999). Further, the findings indicate that even within academic Web spaces the proportion of scientifically reliable publications may be small compared to other kinds of academically associated web contents and may be blurred by other social groups interfering in the production processes on the Web, such as by scientists *in spe* (students at all levels) but also by commercial interests or lay men. Some domains may simply be more inclined to produce or receive more links than other domains or genres on the Web whereby compatibility between domains and genres becomes difficult — just as some scientific disciplines produce many more references turning into citations than others do.

4.2 Web Page Property Analysis

Aside from quality assessments web page analyses mainly deal with their contents and message providing characteristics. According to Cronin and McKim (1996, p. 170) "the Web is reshaping the ways in which scholars communicate with one another. New kinds of scholar and proto–scholar publishing are emerging. Thanks to the Web, work in progress, broadsides, early drafts and refereed articles are now almost immediately sharable … with authors able to choose between narrowcasting and broadcasting. And peer review has emerged from the closet to reveal a spectrum of possibilities…". This belief and vision is indeed reality. Webometric analyses of the nature, such as, genres (or types) and their relationships, structures and content properties of web sites and pages, as well as link structures are important in order to understand the virtual highways and their interconnections. Larson (1996) was one of the first information scientists to perform an exploratory analysis of the intellectual structure of cyberspace. Shortly after, Almind and Ingwersen (1997) applied a variety of bibliometric–like methods to the Nordic portion of the Web in order to observe the kinds of page connections and define the typology of web pages actually found at national Nordic level. The methodology involved stratified sampling of web pages and download for local analysis purposes. The findings revealed that each web page, capable of outlinking, on average provided 9 outlinks – a proportion which nowadays still holds as approximation in the exponentially growing Web space (Björneborn, 2004). The contribution also attempted a comparison between the estimated share of scientific web pages and the distribution found in the citation indexes between the Nordic countries. Clearly, the visibility on the Web was quite different from that displayed in the citation databases. Norway, for instance, was much more visible on the Web than in the printed world at the time of analysis.

Bar-Ilan (2000) and Bar-Ilan and Peritz (2000) studied how a topic like 'informetrics' developed over time on the Web, that is, a kind of issue tracking investigation. They applied search engines, whereas, for instance, Björneborn (2004) applied a web crawler in order to investigate the nature of the academic UK Web space, with special emphasis on transversal links, short cuts between disparate topical clusters on the Web, and small world phenomena. In none of these and other similar studies has been used a complete data set, only more or less systematic or stratified sampling. Most often, the sampling methods, owing to the distortion possibilities discussed above, have been something one might call 'convenience sampling'. This means that since the total population frequently is unknown, unavailable, very large, and its properties inhomogeneous, the sampling must be done on an unsatisfyingly small scale from which generalizations probably always are statistically difficult or invalid. Applications of dedicated Web crawlers are thus an improvement because the harvested data are reusable and analysable locally. This issue is also evident for link analysis work.

5. LINK ANALYSES AND WEB IMPACT FACTOR STUDIES

In his classic webometric article on site inlinks (named 'sitations'), Rousseau (1997) analysed the patterns of distribution of web sites and incoming links. Although Rousseau, like Ingwersen (1998) later, made use of the old unstable version of AltaVista, his study operated with 343 downloaded sites for further analysis, retrieved from a query on the search keys 'informetrics' + 'bibliometrics' + 'scientometrics'. The analyses are thus more independent of the Web engine characteristics and more robust. The analyses showed that the distribution of sites followed the omnipresent Lotka distribution. Similarly, Rousseau demonstrated that the distribution of inlinks to those 343 sites also followed a Lotka distribution. The proportion of selflinks was estimated to 30%.

Since then many other types of link analyses have been performed. Either the investigations make use of predefined search profiles or sets of URLs by means of commercial search engines, or they apply personal crawlers. First, we briefly discuss Web Impact Factor (WIF) analyses, because there are some methodological problems connected to such analyses. This is followed by other selected link analyses with methodological implications.

5.1 Web Impact Factor Analysis

Ingwersen demonstrated (1998) the difference between counts of inlinks and counts of inlinking pages in his attempt to calculate the Web Impact Factors (WIF) for national domains and individual sites[3]. The underlying idea was that the WIF could say something about the awareness, authority or recognition of national sites (on average) or individual sites – but not necessarily quality. The study found three interesting results relevant from a methodological perspective. 1) Since the AltaVista search engine cannot count the actual number of inlinks to particular sites, but only the number of *pages* that are sources of an inlink at least once, selflinking will not influence the overall WIF. The external node inlinks, for instance, site inlinks, or TLD inlinks, hence becomes the important score to observe. This is because for each new web page within a given site providing one or more links to its own site, both the numerator and the denominator increase with the score 'one', given that the analysis unit is web pages. With aggregation into site or higher levels, this phenomenon does not matter. 2) The WIFs for

[3] Note that prior to Ingwersen, Rodriguez i Gairin (1997) had introduced the concept of information impact on the Internet in a Spanish documentation journal.

individual web sites was more unreliable than that of the top-level domains, such as countries. This was, however, owing to the instability of the 'old' AltaVista at that time found later (Rousseau, 1999). 3) The variance in the WIF calculations, also between engines, could be applied as a Web engine *evaluation measure,* i.e., as an indicator of engine performance. However, the instability and variance was probably fortunate, since it already, in the case of Ingwersen (1998), gave cause to prudence in applying the methodology and the interpretation of results.

In connection to the second result in Ingwersen's study, and with the idea of comparisons to, for instance, citation data or other classical parameters in mind, Smith (1999) as well as Thelwall (2000) further investigated the variance phenomena; however, still applying the unstable AltaVista version. Fortunately, exactly owing to the observed variations they both became suspicious about the coverage and retrieval properties of the engine(s). Had the results continuously been stable, etc. during these reproduced experimental trials, one might not necessarily immediately have questioned the methodology.

Smith (1999) demonstrated some periodic and robust data collection methods and showed how results became distorted owing to retrieval of *noise pages*, e.g., Indonesia (domain code: .id) showed very high WIF because of the retrieval of the URL element 'id' in many sites other than Indonesian. He also showed that the longer the URL string searched for, the more reliable the result. The context of the string should assure its uniqueness. However, later unpublished studies of the actual coverage of the engines, including AltaVista, with respect to the known pages and links on our own local server (ax.db.dk) demonstrated that they do not penetrate to all pages and links. Thelwall confirmed (2000) this negative result by applying AltaVista, Hotbot, and Infoseek in his analyses. The coverage is *not random* in such a way that the WIF denominator and numerator are influenced in identical ways. In short, at the present state of search engine coverage and retrieval strategies, "the exiting concept of WIF appears to be a relatively crude instrument in practice" Thelwall (2000, p. 188). Thus far the outcome when applying Web engines seems highly problematic, and, as stated by both Rousseau (1999), Smith (1999), Thelwall (2000), and Björneborn and Ingwersen (2001, 2004), one would have to apply dedicated web crawlers or direct URL searching to download data for local analyses.

Obviously, WIF calculations can be compared to other academic–like impact measures, like classic journal impact factors for journals that are printed and online; personal or institutional citation impact or number of citations received or publications produced; or other economic parameters like IT expenditure or number of staff, etc. A methodologically interesting study in this respect was the use an alternative WIF compared to the

Research Assessment Exercise (RAE) and research productivity and staff size in UK computer science departments by Li et al. (2003). Both AltaVista and a special crawler were used to collect link and page data (not link page numbers but the actual number of links was retrieved). Two kinds of WIFs were calculated: one with staff size per department as denominator and one with department web pages as denominator. The former WIFs correlated significantly with their *RAE ratings* whereas the latter did less well. The numerator values alone, i.e., the number of inlinks to computer departments, correlated significantly with the *research productivity* of the departments. The RAE rating correlation was interesting since Thomas and Willet (2000) did not find significant correlations between inlinks and RAE for UK LIS departments. The staff number per department seems a better indicator of departmental size than web pages numbers. Probably, there are too many pages per department of less research significance.

The major problems with the WIF are its reliability and its interpretation – as for other kinds of scientometric impact factors. The operational variable is well detectable (the links), although less robust than citations, while the theoretical variable, its meaning, is obscure or only understood to a certain degree.

5.2 Other Link Analysis Issues and Link Motivation Studies

Many comparative analyses have been done at a large scale by the Thelwall project team mainly by means of specialised crawlers (see Thelwall, Vaughan and Björneborn (2005) for methodological details). The general trend was found to be that *links* are better sources of information (Oppenheim, 2000) or indicators than web pages at directory and domain levels (Thelwall and Tang, 2003, p. 156).

Wilkinson et al. (2003) used a random sample of 414 links between UK universities. The links were classified according to scholarly content. Less than 1% dealt with contents equivalent to a scholarly article, whereas 90% was created for a variety of academic or scholarly reasons, including teaching. Wilkinson et al. (2003) showed that links to academic sites are not made solely for formal scholarly motivations. Link counts thus measure a host of *informal* scholar communication. This divide of links can be compared to the Jepsen et al. study (2004) of web page content types of academic nature discussed above. Thelwall and Tang (2003, p. 157) outlined a range of link research connected to academic web entities. For instance, interlinking between UK universities decreased with geographic distance. They found that the correlation between link count and research productivity also exist for Taiwan, outside the UK and Australia. Further, “… the most highly targeted pages, at least in the UK, typically have little direct scholar content, e.g., university home pages, and demonstrate clear disciplinary and role biases. For example, a page may be highly targeted because it has an information dissemination purpose, or is related to computing or general university education issues." Evidently, data collection, sampling and analyses must take into account both link and page roles and types or genres which introduces issues and problems of classification and typology, i.e., subjective interpretation.

Link creation motivation studies are central for developing an understanding of how counts of links should be interpreted. Nevertheless, they have tended to lag behind statistical correlation studies. According to Thelwall, Vaughan and Björneborn (2005), most motivation studies have actually investigated the context of links, rather than attempting to directly ascertain author motivations. Motivation studies should be viewed in the context of what is known about *web use* in general. It is important to understand that web use is not determined by technology, it is context–specific (Hine, 2000). In particular, academics use the Web in many different (in)formal ways, and this is likely to continue to be true (Kling and McKim, 2000).

In relation to science and technology studies there are several types of formal academic web communication to investigate. For example, traditional journals in paper format, also available in electronic form; real peer reviewed e-journals; university online series, peer or non-peer reviewed; pre-prints in circulation (forever?) prior to final submission; etc. In all these cases, traditional references and citations *and* outlinks and inlinks are central properties to study. Kim (2000) made, for instance, a detailed investigation into authors' motivations for creating outlinks in e-journal articles. These were found to extend paper citation motivations. New ones were *functional*, that is, relating to accessibility and richness of electronic resources.

One may hypothesize that when moving more into technical/commercial web spaces the more functional and rhetorical are the outlinks. Rhetorical links, as rhetorical references, link to *authoritative* web pages or pages that are *profitable* to link to. For instance, this kind of linking can be done with the purpose of self–presentation and emphasis, showing off professional relationships and collaboration. In addition, such pages may be 'hubs' (Kleinberg, 1999), that is, having many outlinks themselves, also functioning as *web junctions*. Then we move into more functional and navigational linking motives, such as, drawing attention to relevant pages, to share knowledge, experiences, etc.

Within the academic web space one may expect also to find *normative* outlinks, aside from rhetorical and functional ones. For each of these generic types of outlinks there exist a large number of specific reasons for outlinking that are dependent on scientific environments and domains and communication media. Analogously with references, normative linking motives could be acknowledging support, sponsorships, assistance, and providing information of a variety of commercial, academic or entertainment purposes. We have seen above that there exist significant correlations between (normative–like) formal inlinks and research productivity but that functional and rhetoric (informal) linking also display a, albeit weaker, correlation to productivity. Future investigations may reveal associations and degrees of correlations to other interesting parameters significant from the point of view of SandT evaluations. Hence, there seems to be more to distinguishing between inlink genres than commonly done in relation to citation analyses. This seems to be caused by the rather unconventional and fuzzy way linking is done, compared to giving scientific reference on lists, later to be broken up into citations to be counted. Clearly, in traditional citation analysis the motivations for making references may not really matter on large-scale citation analysis because opposite motives become neutralised.

One interesting property of the web linking behavior is that *negative outlinks* are rare or non-existent – in contrast to traditional scientific references. One should also bear in mind the dynamic nature of the Web, i.e., that *time* plays a predominant role. Ageing, i.e., generation, maturity, obsolescence, decline, and death happen faster and are probably less predictive on the Web than in traditional scientific literature (Glänzel, 2003). Methodologically speaking, this dynamic characteristic of the Web makes data collection and analysis highly cumbersome, compared to using the traditional citation databases

6. WEB ENGINE LOG STUDIES OF INTERACTION AND USE

The majority of studies of Web interaction focuses on single sites and is based on server log analyses. This kind of webometrics is a natural bridge to the other major research discipline within information science: integrated information seeking and retrieval studies. For a deeply detailed overview of Web searching research the Jansen and Pooch review (2001) is recommended. The web server log captures ordinary persons' web searching processes and provides data that is useful, not only for web interface and presentation design, but also for the interpretation of the social and psychological impact of the Web on people.

Notwithstanding, there are surprisingly few studies that have focused on *user–centered surveys*, i.e., on the searcher side of Web transactions, e.g., children's and high school students' use of the Web to solve assigned specific search tasks (Ingwersen and Järvelin, (forthcoming).

6.1 Large-Scale Web Engine Studies

Large-scaled Web engine studies are often based on log analysis. The Excite studies reported by Jansen, Spink and Saracevic (2000) and Spink et al. (2001) were preceded by the AltaVista study (Silverstein et al., 1999). Later, Wang, Berry and Yang (2003) reported the longitudinal study of an academic Web server over 4 years, 1997 to 2001.

The major limitations of these studies include that they only catch a narrow facet of searchers' Web interaction. The searcher, his/her intentionality, strategies, and motivations are hardly known. On the other hand, log analysis is an easy way of taking hold of data, which can be treated with quantitative methods. We can use the studies to obtain statistically significant data about user selection of search keys and use of syntax in queries.

Silverstein and colleagues (1999) performed an analysis of approximately 1,000 million requests, or about 575 million non-empty queries, from Alta-Vista. Their findings support the notion that Web users behave differently from users of traditional IR systems, they use few query terms, not the advanced IR features, investigate only a small portion of the result list, and seldom modify queries. It is, however, impossible to tell what the situation would have been like if the search engines had similar response times and the same features that professional IR systems have. Aside from searching for known items by means of URLs it is difficult directly to assess the kind of information needs that underlie the queries posed to the systems. The method used for distinguishing between searchers was a combination of the use of cookies by the searchers and IP addresses. That method is not perfect since cookies can be disabled, different searchers can apply the same browser, and floating IP addresses can be assigned to computers. A method to separate sessions is to define a session, as done by Silverstein et al. (1999) as "a series of queries by a single user made within a small range of time". After 5 minutes of searcher inactivity a session is timed out.

Jansen, Spink, and Saracevic (2000) analysed more than 50,000 queries in the query log provided by the Excite search engine, and probably made from cookies. However, the paper says nothing about whether users search for different topics during a session, i.e., one does not know if they tried to solve more than one task in one session. A follow up study based on analysis of one million queries in Excite (Spink et al., 2001) showed that searchers moved towards even shorter queries and that they viewed fewer pages of results per query.

The third large *longitudinal* investigation by Wang, Berry, and Yang (2003) analysed more than 540,000 user queries submitted to an academic Web server from 1997 to 2001. Their log file and queries did not include IP addresses of individual searchers due to privacy concerns. Hence, the sessions of the individual searchers could not be identified from the log data (p. 744). Nevertheless, the study demonstrates valuable results on query level statistics, which reveal users' search activities, as well as the actual queries that uncover both topics and linguistic structures. The observations reported are thus on the user population as a whole.

6.2 User–Centered Surveys

A different and interesting kind of study, still applying server logs but viewing the processes distinctively from a user oriented point of view, can be found in Catledge and Pitkow (1995). They carried out a longitudinal survey at the Georgia Institute of Technology on 107 persons belonging to the Institute who agreed to have their *client logs* captured over a period of three weeks. The client logs contained the URL of the users' current and target page, as well as information on the technique they used to access the target. The data was more controlled than in the previous studies above and analyzed to compute path lengths and frequency of paths and to distinguish between kinds of web users. The survey also gave some insight into which techniques and tools are being used to browse the Web, e.g., following links and using the back button as means of accessing web pages.

User–centered evaluations and direct observations of human interaction with Web search engines have started to evolve by assessing effectiveness as well as usability factors, such as screen layout, and searcher behavior during interaction. Commonly log protocols are created from monitoring actual searchers' seeking processes, e.g., by means of video or talking aloud recordings, screen and keyboard logging and forms of pre and post interviews. Session length is thus known and often applied as an IR performance parameter. Peoples' own information problems as well as assigned topics or search task situations are used as instigators of the process. See for instance the thematic issue of JASIST (Spink, 2002) for a variety of approaches to this kind of surveys.

7. CONCLUDING REMARKS

The contribution has attempted to demonstrate the relationships between the variety of 'metrical' research areas associated with library and information science, within the framework of its established sub-field informetrics. Fundamentally, *webometrics* is referred to as belonging to cybermetrics and covered by an expanded concept of bibliometrics. Further, the contribution has made an effort to establish a consensus in connection with the outlined link terminology and web node levels of analysis.

The terminology was then applied to discussing methodological facets of webometric research, in particular concerned with academic web spaces, although the methods may very well be applied to other kinds of the Web as well. In relation to the academic part of the publicly available Web an essential aspect was to stress that the analogy between outlinks and inlinks on the one hand, and academic references and citations on the other, should *not* be taken too far. Linking motivations display a greater variety, also on the academic web, than in traditional formal scholar communication. Increasingly, the Web demonstrates tracks of informal interaction and communicative behavior. The most important problematic issues to be aware of when making data collection and analyses for studies of the Web were seen to be:

– Methods of initiating web data collection and their comprehensiveness, e.g., by URLs or appropriate search keys dealing with web page content;
– Search engine and/or specialised crawler retrieval strategies and update frequency of engines;
– Accessibility variations (for download) between search engines;
– Page ranking algorithm applied by the engines used;
– National, genre, and other biases;
– Crawl depth limitations at web sites;
– Page number limitation per site visited;
– Omissions of web pages. Pages are isolated owing to lacking inlinks. The crawlers may not comply with page format, they are protected by measures that do not allow mining or crawling, or by passwords; or servers are momentarily shot down;
– Non-integration of personal home pages in institutional link structures within a site;
– Different domain names used for the same entities under study, e.g., (inter)national corporations under .com, .net, .dk;
– Quality variation as to page genres and other kinds of data on the Web;
– Sampling methods and significance.

Aside from such problematic issues of data isolation, it is important to be aware of *what is measured*. As stated by Björneborn and Ingwersen (2004), there is, for instance, a rather large difference between counting the real number of inlinks to a web site or page and counting the number of in-neighbours in the shape of web pages (or sites) inlinking at least once to some web node. This difference is often overlooked, in both calculus and applying the terminology.

Lastly, the distinction between *web node levels,* its terminological impact, and the application of a consistent diagram notation is necessary if the topology of the Web is to be understood and investigated. There exists a constant possibility of loosing the point of perspective in such analysis, in particular if terminological rigor is lacking.

REFERENCES

Almind, T.C., Ingwersen, P. (1997). Informetric analyses on the World Wide Web: methodological approaches to "webometrics'. *Journal of Documentation*, 53 (4), 404–426.

Allen, E.S., Burke, J.M., Welch, M.E., Rieseberg, L.H. (1999). How reliable is science information on the Web? *Science*, 402, 722.

Bar-Ilan, J. (1997). The 'mad cow disease': Usenet newsgroups and bibliometric laws. *Scientometrics*, 39 (1), 29–55.

Bar-Ilan, J. (1999). Search engine results over time: A case study on search engine stability. Cybermetrics, 2/3 (1999), paper 1. Visited 08.11.2003: http://www.cindoc.csic.es/cybermetrics/articles/v2i1p1.html.

Bar-Ilan, J. (2000). The Web as an information resource on informetrics? A content analysis. *Journal of the American Society for Information Scienc.* 51, 432–443.

Bar-Ilan, J. (2001). Data collection methods on the Web for informetric purposes: A review and analysis. *Scientometrics*, 50 (1), 7–32.

Bar-Ilan, J. (2002). Methods for measuring search engine performance over time. *Journal of the American Society for Information Science and Technology,* 53 (4), 308–319.

Bar-Ilan, J., Peritz, B.C. (2000). The life span of a specific topic on the Web. The case of ' 'informetrics': A quantitative analysis, *Scientometrics.* 46, 371–382.

Björneborn, L. (2001). *Small–world linkage and co-linkage*. Proceedings of the 12th ACM Conference on Hypertext and Hypermedia (pp. 133–134). New York: ACM Press.

Björneborn, L. (2004). *Small–world link structures across an academic Web space: a library and information science approach*. PhD Thesis. Royal School of Library and Information Science, Denmark. http://www.db.dk/dbi/samling/phd/lennartbjoerneborn-phd.pdf.

Björneborn, L., Ingwersen, P. (2001). Perspectives of webometrics. *Scientometrics*, 50, 65–82.

Björneborn, L., Ingwersen, P. (2004). Towards a basic framework for webometrics. *Journal of American Society for Information Science and Technology* (in press).

Brin, S., Page, L. (1998). The anatomy of a large scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30 (1–7), 107–117.

Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., Wiener, J. (2000). Graph structure in the Web. *Computer Networks*, 33 (1–6), 309–320.

Brookes, B.C. (1990). *Biblio-, sciento-, infor-metrics??? What are we talking about?* In: L. Egghe, R. Rousseau (Eds.), Informetrics 89/90: Second International Conference on Bibliometrics, Scientometrics and Informetrics (pp. 31–43). Amsterdam: Elsevier.

Catledge, L. D., Pitkow, J. E. (1995). Characterizing browsing strategies in the World-Wide Web. *Computer Networks and ISDN Systems*, 27 (6), 1065–1073.

Chen, C., Newman, J., Newman, R., Rada, R. (1998). How did university departments interweave the Web: a study of connectivity and underlying factors. *Interacting with Computers,* 10, 353–373.

Clarke, S.J., Willett, P. (1997). Estimating the recall performance of Web search engines. *Aslib Proceedings*, 49, 184–189.

Courtois, M.P., Berry, M.W. (1999). Results ranking in Web search engines, *Online*, May/June, 39–46.

Cronin, B. (2001). Bibliometrics and beyond: some thoughts on web–based citation analysis. *Journal of Information Science*, 27 (1), 1–7.

Cronin, B., McKim, G. (1996). Science and scholarship on the World Wide Web: A North American perspective. *Journal of Documentation*, 52, 163–172.

Cui, L. (1999). Rating health Web sites using the principles of citation analysis: A bibliometric approach. *Journal of Medical Internet Research*, 1 (1), e4 (ISSN: 1438–8871). Visited 08.11.2003: http://www.jmir.org/1999/1/e4/index.htm.

Egghe, L., Rousseau, R. (1990). *Introduction to informetrics: quantitative methods in library, documentation and information science.* Amsterdam: Elsevier.

Glänzel, W. (2003). *Personal communication*. Available – visited 08.11.2003. http://www.oud.niwi.knaw.nl/nerdi/lectures/glanzel.pdf

Herring, S.C. (2002). Computer–mediated communication on the Internet. *Annual Review of Information Science and Technology*, 36, 109–168.

Hine, C. (2000). *Virtual Ethnography*. London: Sage.

Henzinger, M.R., Heydon, A., Mitzenmacher, M., Najork, M. (2000). *On near–uniform URL sampling*. Proceedings of the 9th International World Wide Web Conference, May 2000. Computer Networks, 33 (1–6), 295–308.

Hou, J.Y. & Zhang, Y. (2003). Effectively finding relevant Web pages from linkage information. *IEEE Transactions on Knowledge and Data Engineering*, 15 (4), 940–951.

Ingwersen, P. (1998). The calculation of Web Impact Factors. *Journal of Documentation*, 54, 236–243.

Ingwersen, P., Järvelin, K. *The Turn: integration of information seeking and retrieval in context*. Kluver (forthcoming).

Jansen, B.J., Pooch, U. (2001). A review of Web searching studies and a framework for future research. *Journal of the American Society for Information Science*, 52 (3), 235–246.

Jansen, B. J., Spink, A., Saracevic, T. (2000). Real life, real users, and real needs: a study and analysis of user queries on the web. *Information Processing & Management*, 36, 207–227.

Jepsen, E.T., Seiden, P., Ingwersen, P., Björneborn, L., Borlund, P. (2004). Characteristics of scientific Web publications: Preliminary data gathering and analysis. *Journal of American Society for Information Science and Technology* (in press).

Kleinberg, J.M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46 (5), 604–632.

Kleinberg, J., Kumar, R., Raghavan, P., Rajagopalan, S., Tomkins, A. (1999). The Web as a graph: measurements, models and methods. *Lecture Notes in Computer Sc.*, 1627, 1–18.

Kling, R., McKim, G. (2000). Not just a matter of time: field differences in the shaping of electronic media in supporting scientific communication. *Journal of the American Society for Information Science*, 51 (14), 1306–1320.

Larson, R. (1996). *Bibliometrics of the World Wide Web: an exploratory analysis of the intellectual structure of cyberspace*. Proceedings of the 59th Annual Meeting of the American Society for Information Science, 33, 71–78.

Lawrence, S., Giles, C.L. (1998). Searching the World Wide Web. *Science*, 280, 98–100.

Lawrence, S., Giles, C. L. (1999). Accessibility and distribution of information on the Web. *Nature*, 400, 107–110.

Li, X.M., Thelwall, M., Musgrove, P., Wilkinson, D. (2003). The relationship between the WIFs or inlinks of Computer Science Departments in UK and their RAE ratings or research productivities in 2001. *Scientometrics*, 57 (2), 239–255.

Molyneux, R.E., Williams, R.V. (1999). Measuring the Internet. *Annual Review of Information Science and Technology*, 34, 287–339.

Oppenheim, C., Morris, A., McKnight, C. (2000). The evaluation of WWW search engines. *Journal of Documentation*, 56, 190–211.

Otte, E., Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*, 28 (6), 441–454.

Park, H.W., Thelwall, M. (2003). Hyperlink analyses of the World Wide Web: A review. Journal of Computer–Mediated Communication, 8 (4). Visited 08.11.2003: http://www.ascusc.org/jcmc/vol8/issue4/park.html.

Pinski, G., Narin, F. (1976). Citation influences for journal aggregates of scientific publications: theory, with applications to the literature of physics. Information Processing and Management, 12, 297–312.

Pirolli, P., Pitkow, J., Rao, R. (1996). Silk from a sow's ear: extracting usable structures from the Web. CHI 96 Electronic Proceedings. Visited 08.11.2003: http://www.acm.org/sigchi/chi96/proceedings/papers/Pirolli_2/pp2.html.

Rodriguez I Gairin, J.M. (1997). Volorando el impacto de la informacion en Internet: Altavista, el "Citation Index' de la Red. Revista Espanola de Documentacion Scientifica, 20 (2), 175–181.

Rousseau, R. (1997). Sitations: an exploratory study. Cybermetrics, 1 (1). Visited 08.11.2003: http://www.cindoc.csic.es/cybermetrics/articles/v1i1p1.html.

Rousseau, R. (1999). Daily time series of common single word searches in AltaVista and NorthernLight. Cybermetrics, 2/3, paper 2. Visited 08.11.2003: http://www.cindoc.csic.es/cybermetrics/articles/v2i1p2.html.

Rousseau, R. (2001). *Evolution in time of the number of hits in keyword searches on the Internet during one year, with special attention to the use of the word euro.* In M. Davis, C. Wilson (Eds.), Proc. of the 8th Int. Conf. on Scientometrics & Informetrics. Sydney, 619–627.

Rusmevichientong, P., Pennock, D.M., Lawrence, S., Giles, S.L. (2001). *Methods for sampling pages uniformly from the Web*. In Proceedings of the AAAI Fall Symposium on Using Uncertainty within Computation, 121–128.

Silverstein, C., Henzinger, M., Marais, H., Moricz, M. (1999). Analysis of a very large Web search engine query log. *SIGIR Forum*, 33 (1): 6–12.

Smith, A.G. (1999). A tale of two web spaces: comparing sites using web impact factors. *Journal of Documentation*, 55, 577–592.

Spink, A. (2002). Introduction to the special issue on Web research. *Journal of the American Society for Information Science & Technology*, 53 (2), 65–66.

Spink, A., Wolfram, D, Jansen, B. J., Saracevic, T. (2001). Searching the Web: the public and their queries. *Journal of the American Society for Information Science*, 52 (3), 226–234.

Snyder, H., Rosenbaum, H. (1999). Can search engines be used as tools for web-link analysis? A critical view. *Journal of Documentation*, 55, 375–384.

Tague-Sutcliffe, J. (1992). An introduction to informetrics. *Information Processing & Management*, 28 (1), 1–3.

Thelwall, M. (2000). Web impact factors and search engine coverage. *Journal of Documentation*, 56, 185–189.

Thelwall, M. (2001a). Extracting macroscopic information from web links. *Journal of the American Society for Information Science and Technology,* 52 (13), 1157–1168.

Thelwall, M. (2001b). The responsiveness of search engine indexes, *Cybermetrics*, 5 (1). Visited 08.11.2003: http://www.cindoc.csic.es/cybermetrics/articles/v5i1p1.html.

Thelwall, M. (2001c) A Web crawler design for data mining. *Journal of Information Science,* 27 (5), 319–325.

Thelwall, M., Tang, R. (2003). Disciplinary and linguistic considerations for academic Web linking: an exploratory hyperlink mediated study with Mainland China and Taiwan. *Scientometrics*, 58 (1), 155–181.

Thelwall, M., Vaughan, L., Björneborn, L. (2005). Webometrics. *Annual Review of Information Science and Technology*, 39 (in press).

Thomas, O., Willett, P. (2000). Webometric analysis of departments of librarianship and information science. *Journal of Information Science*, 26 (6), 421–428.

Vaughan, L., Thelwall, M. (2003). Scholarly use of the Web: What are the key inducers of links to journal web sites? *Journal of the American Society for Information Science and Technology*, 54 (1), 29–38.

Vaughan, L., Thelwall, M. (2004). Search engine coverage bias: evidence and possible causes. *Information Processing & Management* (to appear).

Wilkinson, D., Harries, G., Thelwall, M., Price, E. (2003). Motivations for academic Web site interlinking: evidence for the Web as a novel source of information on information scholarly communication. *Journal of Information Science*, 29 (1), 59–66.

Chapter 16

DESCRIPTIVE VERSUS EVALUATIVE BIBLIOMETRICS

Monitoring and Assessing of National R&D Systems

Thed van Leeuwen

Centre for Science and Technology Studies (CWTS), Leiden University, the Netherlands
E-mail: leeuwen@cwts.leidenuniv.nl

Abstract: This paper covers the differences between two separate bibliometric approaches, labelled 'descriptive' versus 'evaluative', or top down versus bottom up. The most important difference between these two approaches is found in the level of validity of the underlying research output. Whilst the publications in a top down approach, having a descriptive character, are collected by following general characteristics of these publications (such as country names, or fields), the consequence is that findings from such studies have a 'meaning' that is limited with respect to actual research assessment. On the other hand, in a bottom up approach the publications are collected from individual oeuvres of scientists, including a process of verification by the researchers involved. This procedure contributes significantly to the validity of the publication material, and consequently research assessment procedures can be based on the results of this type of bibliometric analyses. A strong focus in the paper will be on the actual application of bibliometric analysis within research assessment procedures, in particular within the UK and the Netherlands.

1. INTRODUCTION

The last decades have shown a steady increase in the growth of assessments of science, at different levels in the science system, and by different actors within the science system. Whilst peer review was the long existing practice for assessing science and its actors, the last decade has shown a strong increase of the application of quantitative techniques in the 'assessment' of science. As a result of the development of the field of quantitative studies of science and technology, and the growing awareness in

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 373-388.

the western world of the necessity of the science system periodically to provide some sort of accountability of science and scientists to society, we observed a development in which modern day science is both monitored as well as assessed on a regular basis. In a number of countries (such as France and the Netherlands) national facilities were created to monitor the national science system, in an international perspective, whilst in a number of countries actual assessments at a national level of the science system was initiated (e.g., the UK, and the Netherlands). In this paper, we will discuss the various approaches in the monitoring and assessing on both the international and national level, as well the differences between actual assessments within the UK and the Netherlands (an excellent overview in a more general sense has recently been given by Geuna and Martin, 2003). The difference between both methods evolves around the distinction between descriptive and evaluative bibliometrics.

2. DESCRIPTIVE AND EVALUATIVE BIBLIOMETRICS

Bibliometrics is the field of science that deals with the development and application of quantitative measures and indicators for sciences and technology, based on bibliographic information. This bibliographic information is the representation of codified knowledge as can be found in a diversity of scientific output types, such as serial literature, books, and book chapters, conference proceedings, patents, etc. And although the use of serial literature is not evenly distributed over fields of science (for instance, the dominant use of conference literature within some fields of the technical sciences), bibliometric studies start from the assumption that the most important findings of scientific research finally end up in the international serial literature. This, however, means that, in general, bibliometrics is less applicable in those fields of science in which the internationally oriented scientific journal is not the main medium for communicating research findings to the (international) community in those fields. Next to this application in the analysis and assessment of the development of science, bibliometrics focuses on the development of technology as well. Here patents are the main source of analysis.

Evaluative bibliometrics, as introduced by Narin (1976), is the application of bibliometrics which focuses particularly on the evaluation of scientific activity, and more, in particular, on quality aspects of scientific performance. In general, evaluation in itself is focused on the control of quality, so that, more specifically, research evaluation is focusing on the safeguarding of scientific quality. Scientific quality is a rather diverse concept, and a synonym for several different meanings. From a bibliometricians point of view, scientific quality is the synonym for scientific merit (Moed, 1989), representing scientific influence, particularly (international) scientific visibility. Quantitatively, it is based on both scientific output and impact measurement. The circumstance that distinguishes evaluative bibliometrics from descriptive bibliometrics is the degree of validity and reliability of the publication data, underlying a bibliometric analysis. Whilst a focus on a whole country or even a whole university normally uses a so called top down approach, in which all output is collected using the address information of the publications in the ISI citation indexes, the sight on smaller, lower level organisational structures remains clouded. Therefore this approach only offers insight on rather high levels of aggregation, and does not allow for any detailed conclusions within a country or organisation. On the other hand, in a so called bottom up approach in which specific target groups are assessed, these groups should be asked to co-operate in collecting publication data, thereby seriously and significantly contributing to the validity and reliability of the resulting output and impact scores.

An important issue with respect to the distinction between evaluative and descriptive relates to the interpretation of bibliometric results from both approaches. Whilst it is clear to both the bibliometrician, as well as to the direct user of the results of a top down or descriptive study, to what extent conclusions can be drawn from the results of such a study, other users who are not very well versed in the ins and outs of bibliometrics in general or that specific study might overlook these limitations. Comparisons made at higher levels of aggregation still do not allow for conclusions at lower levels of aggregation. Therefore comparisons made in top down studies do not have the same meaning as the results from a bottom up study.

3. MONITORING SCIENCE THOUGH MACRO-BIBLIOMETRICS

In recent years bibliometrics in its broadest sense has become more and more important at the national or supra-national level, as can be concluded from the studies initiated by government agencies in a number of countries and the European Commission. Next to the long time existence of the National Science Foundation (NSF) in the USA, with its periodical reports containing analyses of the American science system, the early 1990s in Europe was a period in which national facilities were created to monitor the national science system. In France the national Observatory of Science and Technology (OST, Observatoire de Science et de Technologie) was founded in Paris in 1990, while in the Netherlands a similar, but more virtual, observatory of science and technology developments was created in 1992 (NOWT, which stands for the Dutch Observatory of Science and Technology). Next to bibliometric indicators of scientific and technological performance, both organisations also provided in their periodic reports other indicators of the science system they were describing, e.g., data on the financing of the national science system, knowledge workers, graduation figures, etc. The reports of both observatories have a strong focus on benchmarking their national system with other countries, and present analyses of regions within a country and across Europe, of societal sectors, and of actors (at the level of organisations) within the national science system. The existence of institutionalised (France) or 'virtual' (the Netherlands) observatories of the development of science and technology clearly indicates that bibliometrics can and does contribute to science policy. This development in France and the Netherlands was followed by the appearance of the publication by the European Commission of a series of European Science and Technology Indicators Reports, of which the third edition appeared in 2003. While the OST report was written in French and the NOWT report in Dutch, the European Report, with its much broader scope (in terms of the geographic area covered), is an English language report, and as such it is more accessible than its French or Dutch counterparts.

Owing to the interest of these agencies, bibliometric studies often have a comparative character, and compare countries, geographical regions, and, in the last couple of years, even universities. As these studies start from a macro level of aggregation, based on the names of countries, cities, postal codes, and universities from the address information attached to publications in the publications retrieved from the ISI databases, their outcomes often lack accuracy. Therefore these studies can only remain comparative and descriptive.

The type of bibliometric description supplied by this kind of study provides insight on a geographical level as well as a cognitive level, that is, at the level of major fields or disciplines of science. Whilst the geographical level is 'distilled' from the address information available in scientific publications, the level of research fields or disciplines is made available by linking the journals in which the publications occurred to research fields and disciplines. The data available for this enrichment is found in the Journal Subject Categories supplied by ISI for journals covered in their databases. Although this journal classification system is far from perfect, and subject to debate within the bibliometric community, it is currently the only system available for bibliometricians, which fits the multidisciplinary character of the ISI citation indexes best.

From the above we see that by linking journal publications to either (aggregated) address information or fields of science or disciplines the macro-level becomes two-dimensional: countries, regions or organisations versus fields or disciplines. However, one still needs to take care with respect to the 'reach' of these analyses. Whilst it is previously argued that some areas of science are far less well represented by the publications covered in the ISI databases (see for example Hicks (1999)), the problem of sufficient representation stretches to the bibliometric analysis at a macro-level as well. In general we find less international journal publication output for scholars in the arts and humanities, and to a lesser extent for those working in some fields of the social and behavioural sciences. Next, we find a somewhat stronger presence in these international ISI–covered journals by scientists from the Anglo Saxon world. For quite a large group of countries, publishing in English in the arts and humanities is less obvious. For instance, results of the research into Spanish culture and history and its influence on Latin America, research focusing on Italian history and archaeology, German literature and French philosophy, to mention only a few areas, appear in journals containing articles written in these respective languages. Only some of these non-English language journals are actually covered in the ISI citations indexes. This does not mean that these research fields are not internationally oriented, but rather reflect the situation that in some areas of scholarship, English is not necessarily the *lingua franca* of current day developments. Contrary to, for example, the Dutch and Scandinavian language areas, the Spanish, French, German end Italian language areas are large enough to allow scientific publishers to operate in a market that is economically sound enough to publish journals in these languages. Remarkably enough, the issue of the language of publication is not limited to the scholars working in the arts and humanities. A study for the German government showed the effect of language of publication in the fields of the medical sciences (Tijssen, van Leeuwen, and van Raan, 2002; van Leeuwen et al., 2000; and van Leeuwen et al., 2001). The discussion of publication language affected heavily the discussion on the value of journal impact measures within the German language scientific arena (Herfarth and Schurmann, 1996; Haller, Hepp and Reinhold, 1997; Rempen, 1998; Beller, 1999; and Kindermann, 1999).

As mentioned above, the macro level approach stretches out to the level of organisations. However, this is not an easy task. Whilst it often seems pretty obvious from the addresses attached to the publications in the ISI databases, which institutes or organisations are meant, many organisational links underlying the published address, indicating the actual structure of an university (e.g., the affiliated academic hospitals, related research institutes, etc) are not clearly visible, and much efforts should be put into cleaning and unifying address information on scientific publications. On the level of macro bibliometrics, simply not enough noise is filtered out to allow for far reaching conclusions at the level of the organisational level, let alone the level **below** the main organisational level.

4. RESEARCH EVALUATION EXERCISES IN THE UK

In the UK scientific research has been monitored on a large scale from 1986 onwards. With the Research Assessment Exercises (RAE), the British Higher Education Funding Councils assess periodically the research performance of British universities. The main purpose of the RAE is 'not just to enable funding to be allocated selectively but also to promote high quality: research in higher education institutions conducting the best research receive the largest portion of the grant' (RAE 2001). As a secondary effect, the results of the RAE inform other funding bodies in the UK and abroad, as the results of these assessments constitute a basis for science policy and strategy decisions. These assessments take place across so called 'units of assessment' (in 2001, 68 of these units were applied), which embrace research activities in broad scientific disciplines such as chemistry, sociology, etc. The grouping of research fields is dynamic, and is established in consultation with the higher education sector itself. The information playing a role in the assessment has developed over time. Whilst the first exercises consisted of mainly publication output assessment, in later exercises additional information was provide for the peer panel. This additional information is provided by the research groups and departments themselves, and includes the number of staff involved, a description of the research conducted and future plans, the funding received, and a short survey of various types of research output. In the assessments peer committees determine scores for university groups and departments, based on the input provided by the submitting groups and departments. As a consequence of the time consuming process, and the high costs involved with the RAE, the discussion on applying bibliometric analyses in the RAE has started. Probably a more important role in future RAEs is played by citation analysis, which has caused a vivid debate on the applicability and validity of citation analysis on the level of the individual researcher (e.g., Warner, 2000a, 2000b), and the undesirable effects citation analysis might have on the behaviour of scientists (Warner, 2003). Whilst the proponents of the application of citation analysis in the research assessment indicate that there exists a strong correlation between RAE scores and impact scores (Oppenheim, 2000; and Norris and Oppenheim, 2003), the (small scale) citation analyses so far applied have certain weaknesses. These will be discussed below. A very thorough, but more general, critique of the RAE is given in a report by PREST, of the University of Manchester (PREST, 2000). Their criticism focuses on specific aspects of the UK situation, e.g., the funding structure or the composition of research assessment, no longer following disciplinary borders, but rather problem and application oriented assessment of research.

The most recent plans of the Higher Education Funding Council for England (HEFCE) intend to strongly reward 'world class' research. This might lead to a further undermining of the research system in the UK, simply because not enough money becomes available to keep research groups and departments at a viable level. The idea is to focus on excellent research, and support, if possible, for the other research. This might lead to concentration of research funding, which in its turn leads to a cyclic process in which the stronger ('funded') groups and institutes become stronger, and the weaker might eventually disappear.

5. PAST EXPERIENCE IN THE NETHERLANDS

In the Netherlands large scale research assessment procedures were initiated by the VSNU (Association of Universities in the Netherlands) in almost all disciplines of the sciences, social sciences and arts and humanities, according to clearly described procedures from a protocol (VSNU, 'Assessment of Research Quality', 1998). This protocol was in principle applicable to all fields of science and scholarship. However, the main characteristic of all the research assessment procedures in which evaluative bibliometrics was applied, was the choice of the research field underlying the assessments. For instance, in chemistry several bibliometric analyses were applied, and in cultural anthropology none. Another common feature of the research evaluations by VSNU was a focus on the research group level. In all these VSNU research evaluation assessments bibliometrics was used as an instrumental supporting tool for the peer review process. An international peer review committee was provided with both qualitative and quantitative information. The qualitative part included a report written by the groups themselves in which an overview was given of the scope of the research involved, publication output, ability to raise funds, whereas the quantitative part comprised a bibliometric report, focusing on the research group level. The committee judges each group on four aspects of their scientific performance (*quality*, *productivity*, *relevance*, and *viability*) on a five-point scale.

Ranging from astronomy and astrophysics, physics, chemistry, biology, the life sciences, electrical engineering, and psychology, bibliometrics was applied in addition to peer review. In the latter field, particularly within the theoretical psychology community, a discussion started about the validity and reliability of bibliometric indicators within the field.

Within these disciplinary research evaluation procedures initiated by the VSNU, bibliometric procedures have been applied in an open, transparent way. Scientists are asked to control and verify journal publications identified in the citation index based data system, and to comment on the results of the bibliometric data collection. The results of this process are discussed with the peer review committee (or its chairman) and are finally reported to the committee. The peer committee has meetings with the leaders of all university research groups involved in a research performance procedure, giving the research group leaders the opportunity to comment on the findings in the bibliometric procedure.

6. FUTURE PROSPECTS FOR EVALUATION IN THE NETHERLANDS

The research performance assessments initiated by VSNU had a nation–wide disciplinary and cyclic (each five or six years) character, focusing on research groups within all universities. Each research group had to provide its own self–evaluation study, containing next to a survey of input data, or personnel, a qualitative description of the research conducted and of the future plans and prospects. This created a considerable 'evaluation bureaucracy' because research groups had to provide much paperwork, similar to the information they had to provide in the paperwork requested by granting organisations (such as national research organisations) but, of course, always in another format, for another time period, etc.

This situation stimulated the argument to revise the evaluation procedure, in such a way that many of the efforts by scientists and science managers within universities would serve both purposes.

While the research performance procedures initiated by VSNU were based on a protocol from 1998, the assessment of publicly funded research is now subject to a new protocol, designed by VSNU, NWO (the national research council, funding research in various disciplines on a proposal basis) and the KNAW (the Royal Dutch Academy of Sciences) (Standard Evaluation protocol (2003 – 2009) for Public Research organisations, VSNU, NWO, and KNAW, January 2003). The purpose of this extension is to evaluate, next to academic research, the research done on behalf of NWO and KNAW (both their funded research, and the research conducted in their research institutes). The new protocol is strongly influenced by a report from the Commission chaired by professor van Bemmel ('Kwaliteit verplicht', 2000).

As mentioned above, the newly applied protocol covers the full range of publicly financed scientific research. Whilst the main goal of the previous system was to monitor the quality of national academic scientific research in comparative perspective, the new system adds another goal to the research evaluation procedures. This goal relates to the management of scientific research itself, within the organisational structure in which it is carried out in. Another important difference between both procedures is the level of focus: the new protocol is more flexible in such a way that it also allows evaluation at levels other than the research group level. As a consequence the initiative and responsibility for research assessment procedures has shifted from the different Chambers (national disciplinary 'boards') within the VSNU organisational structure, to the top of the organisations (universities) subject to research assessment.

A critique which regularly arose in the previous evaluation procedure dealt with the composition of the review committee. It was either too small (in number) or it lacked knowledge about a specific (sub)field. The new procedure allows a more flexible composition of the peer review committee, taking into consideration the specific type of research conducted in a specific organisation. How this will work out in practice is not yet clear, owing to the recent nature of these developments.

The new evaluation procedure should (try to) focus more on the future perspectives of a specific research unit, thereby replacing the previous main characteristic, namely 'accountability', which leads by nature to a more 'historical' result. A perspective on future developments should be included in the self–evaluation study that is obligatory in each evaluation procedure. Interestingly enough, within bibliometric analyses such a distinction between 'looking farther back in the past', and a 'future perspective' has been long included, and is part of the CWTS methodology applied in national assessment procedures under the previous protocol.

In the report by the van Bemmel commission quantitative studies supporting the evaluation procedure have hardly any priority. The term 'citation analysis' is mentioned only once ('Kwaliteit verplicht', pag. 25), thereby indicating the very modest position of quantitative measures in the new evaluation procedure. However, some of the new goals (for instance, a future perspective of the research in a research group) are hard to grasp by means of only qualitative information. Bibliometric analysis has at least the advantage that it can, based on recent past performance, give some objective insight into future perspectives, starting from the hypothesis that results in the recent past are the best predictors for the near future. This raises the question of the competence of peer committees, and the level of the qualitative information provided in the evaluation process, in the sight of the defined goals in future research assessment procedures

Ironically enough, the 2003 Science Budget of the Netherlands indicates that the current research assessment protocol finds little enthusiasm within cycles of policy makers, especially because the national perspective is abandoned within the current approach, an aspect that was already satisfactorily covered in the 1998 Research Assessment Protocol. So little can be said about the direction in which research assessments will be heading in the very near future in the Netherlands.

7. SYNTHESIS

As stated above, the approaches in the UK and the Netherlands with respect to research assessment were different, and especially in the application of bibliometric indicators, some remarkable differences are observed. Whilst the RAE is bibliometrically based on publication assessment, as can be concluded from the statement that 'All forms of research output (books, papers, journals, recordings, performances) are treated equally' (HERO, 2002, Guide to the 2001 Research Assessment Exercise), the research assessments in the Netherlands were, if applicable in the 'field under assessment', full scale analyses of both the output and the impact of research groups, but bibliometrically restricted to journal publications because they can be found in the citation indexes from ISI.

A very important difference between the actual research assessments in both the UK and the Netherlands is found in the relation between the assessment and research funding. Whilst one could argue that this is beneficial for 'the best scientists', the position and application of quantitative techniques in the process of assessment are of major importance. Given the statement above, the question arises of how peer panels are capable of assessing the four items of research output, in an international context, and of how they treat different research output types equally. Here citation analysis might be helpful, but then other problems occur. In a citation analysis of research output that is supposed to treat different output types equally, citation analysis might be helpful to analyse internationally refereed journal publications as covered by the citation indexes of the ISI adequately, but immediately raises the issue of the limitations of citations analysis of the other types of research output mentioned above. As the ISI citation indexes give, in general, an excellent survey of the most important serial literature, the degree with which citations to non-serial literature are representative for the actual impact of these types of sources, remains to be seen. As one might expect, these types of sources have a citation cycle of its own, with a stronger focus on the same type.

Within the current constraints of bibliometric analysis, which are predominantly indicated by the boundaries of the ISI citation indexes, with their focus on journal publications, a number of guidelines can be determined for research evaluation assessment procedures. If the results of the research assessments, expressed in peer panel judgements, are going to form the basis for the funding, and, based on bibliometric data, the applicability of quantitative techniques on the assessed field should be assessed first. In other words, if the types of research output cannot be measured adequately with bibliometric techniques, owing to the nature of the field under assessment, current existing bibliometric techniques hardly contribute to the final assessment of the research conducted in this field. Only in the last couple of years have systematic attempts been made to develop quantitative indicators that can be used within the social sciences and the humanities, in close consultation with scientists in those fields (Luwel et al., 1999). However, this is not yet as developed as the quantitative techniques applied in the natural, life and technical sciences.

Once you have established the applicability of bibliometric techniques in a certain field, a next guideline would be the level of bibliometric analysis. Whilst some proponents of bibliometric techniques in the British RAE propose citation analysis of individual output, which can be aggregated to group scores (Norris and Oppenheim, 2003), a direct group approach involving the complete journal publication of a research group or research department would be more fruitful. Instead of a selection of only four research outputs, a systematic analysis of the complete oeuvre of a group would provide a more thorough insight into the volume and development of a research group, in terms of its research output in journal publications. And although many scientists in the Western world nowadays have access to the ISI citation indexes (mainly through the *Web of Science*), and therefore are able to estimate which of their papers belong to their most highly cited publications, the number of only four publications is a very small number to base a research assessment on, which determines research funding for the upcoming period. The usage of a somewhat larger number of publications, which follows nearly automatically from an approach in which the research group is the focal point, is less sensitive to statistical problems which occur when the bibliometrician focuses on the (small) output of any individual researcher. Another problem is the overlap that can occur between the research outputs of staff members of one group. This might lead to unrealistically high citation scores.

Once the decision is made to apply the bibliometric analysis to only those fields in which bibliometrics can be applied adequately, and to focus the research assessment on the research output at any level above the individual researcher's level, one needs to take care with respect to the data collection procedure. In the data collection phase the distinction between descriptive and evaluative bibliometrics can be expressed most clearly: with the implementation of a verification process, in which the researchers under assessment check and control the collected publication output, the necessary reliability and validity are added to the process. This step makes the whole process more transparent, and provides a certain space in which scientists can argue about the process, add and/or delete publications to their oeuvres, and, in general provide valuable background information on the publications characteristics of a field. Furthermore, this step allows the users of the outcomes of such bibliometric analyses to draw conclusions from the material, that were not possible if the publications were collected 'from a much larger distance' between assessed research field and bibliometricians.

Finally, if the publication data are collected in such a way that assessments can be carried out, perhaps the most important part of the whole process starts, namely, the calculation of bibliometric indicators. Whilst the RAE impact analyses are based on crude citation counts, citation per publication counts, and the number of papers amongst the most highly cited publications in a field, more sophisticated indicators should be calculated. For example, normalised indicators expressing the ratio between the actual impact and expected values are strong indicators (Moed, de Bruin, and van Leeuwen, 1995), but also indicators describing aspects of scientific publishing like the percentage of self-citations, or the percentage of publications not cited can contribute to a more balanced judgement of the research performance of a group.

8. CONCLUSIONS AND DISCUSSION

This paper focused on the distinction between descriptive bibliometrics, as resulting from top down analyses, and applied in national facilities monitoring national R&D systems from a somewhat larger distance, the results of which are normally publicly available, and evaluative bibliometrics, resulting from bottom-up approaches, applied in (national) research assessment procedures of disciplines and research organisations such as universities or research institutes. Here the results of such exercises are mostly kept confidential. For the former type of bibliometric analyses peer information is hardly used, consequently indicating the larger distance of these analyses to the scientific research itself. In the latter type, peer judgements are a necessary ingredient of research assessment. There is a long-standing, ongoing discussion on advantages and disadvantages of peer review in general (a very good survey is given by Nederhof (1988), and by van Raan (1996). The most important advantages relate to the self–organising principles of science as a quality — and particularly reputation — oriented community, in combination with a reasonable degree of consensus about the direction in which scientific developments should/could proceed. On the other hand, important drawbacks of peer review relate to the composition of the review committee: aspects of subjectivity, conflicts of interest, high costs involved, and insufficient recognition of young promising scientists and/or recent promising scientific developments by the members of a peer review committee. And whilst the application of bibliometric techniques in assessment procedures will not solve all these problems or disadvantages, there most certainly can be a future for the application of bibliometrics in research assessment procedures. This requires a stronger consensus amongst the people in the bibliometric community, a better presentation of the indicators of use in those processes, and also a testing of the robustness of indicators applied in research performance assessment exercises (van Leeuwen et al., 2003).

An example of the difficulties for bibliometricians, let alone laymen, to fully understand the handling of underlying data in the construction of indicators can be found in the calculation of the Impact Factor. Although defined quite clearly, the actual reconstruction of the composing parts for interested users is quite difficult. In an article in *Nature* in 2002 by Moed, the limitations of journal impact factors for the bibliometric practice were shown. However, some of the arguments which were addressed against the application of journal impact factors in the bibliometric practice have a farther reaching meaning, which goes beyond the application of only Impact Factors. The application of bibliometrics, in general, within a research performance measurement procedure requires answers to major questions. For the bibliometrician the first question relates to the issue of visibility of its subject within the international scientific literature, in other words, to what extent is the application of bibliometrics (based on bibliographic data in the ISI databases) focusing on a substantial part of the research output of the subject of study? Normally a bibliometric analysis contains both *output* and *impact* indicators, and the latter can be calculated in a reliable manner only on the basis of the citation index data. These indexes of the ISI are, up until now, the only multi-disciplinary databases containing the complete reference lists to previous scientific journal literature. This allows the bibliometrician to calculate impact indicators, in which the visibility or influence of these previous scientific publications can be expressed. So a first issue that needs to be addressed before starting a bibliometric study relates to the focus on the (English language) international journal literature, and its coverage within the databases of ISI.

If this first issue is answered in such a way that bibliometric data can be collected, the bibliometrician should ask him/herself the question: 'What percentage of the publications of a research group is indeed covered in the international journal literature, and to what extent is this influenced by the specific characteristics of either the (sub)field concerned, or of the research group?' Within a large discipline, such as e.g., chemistry, differences may exist within that discipline. For instance, in an evaluation of chemistry as a whole, the typical 'classical' fields of chemistry such as analytical chemistry are analysed in combination with chemical engineering, biotechnology, and biochemistry. We are then confronted with considerable diversity in publication practices amongst these fields within a discipline, such as a stronger focus on conference proceedings, higher publication rates, a more 'dense' citation traffic, etc. Thus, the intra-disciplinary variation can be as broad as the interdisciplinary or intra-science variation as a whole.

Once these topics are adequately dealt with, a bibliometric analysis should comprise more than only crude publication counts, citation counts, or mean citations per publication ratios, exactly because these indicators do not indicate the performance of a research groups in such way that it can be compared to other units within the assessment, e.g., how does a group in biology then compare with a group in the social sciences? Therefore the bibliometric community should stress the necessity of a certain standardisation in which normalised indicators play a central role. This is a general improvement; other, more conceptual, improvements could be found in a broadening of the scope of bibliometric analyses, in which other aspects of the performance of a research group or institute become visible. One of such techniques can be found in the so called research profiles, in which a spectral analysis is given of the distribution of the output of a group over research fields, in combination with impact scores. In the current bibliometric practice of CWTS these profiles are generated on the basis of the ISI Journal Subject Categories, but this information can be replaced by other information indicating the research topics of a certain field. These profiles then indicate the multi-disciplinary character of research groups from one group to another, and as such indicate a certain resemblance or difference between groups in a research assessment process.

Another useful instrument is formed by scientific cooperation profiles. These profiles, based on the addresses attached to a publication in the citation indexes, indicate the different types of orientation research groups can have in the world outside their own 'environment', and especially indicate the international scope of both the group and the research field as a whole, by comparing scientific cooperation profiles among groups. These analyses could be extended by network analyses, indicating the partners of research groups, and the success of this cooperation.

Finally, the assessment of research groups performance can be extended with analyses of the 'reception' of a groups output. Here the focus can be on who is citing your recent research output, in which journals, and to what extent do we see knowledge transfer between fields.

REFERENCES

Beller, F.K. (1999). Der Zusammenhang zwischen Index Medicus, dem Impact Factor und der deutschen Sprache. *Geburtshilfe und Frauenheilkunde*, 59, 53–56.

Geuna, A., Martin, B.R. (2003). University Research Evaluation and funding: an international comparison. *Minerva*, 41, 277–304.

Haller, U., Hepp, H., Reinhold, E. (1997). Tötet der 'Impact Factor' die deutsche Sprache ? *Gynakologisch–geburtshilfliche Rundschau*, 37, 117–118.

Herfarth, C., Schurmann, G., (1996). Deutsche klinische Zeitschriften und der Impact Factor. *Chirurg* 67, 297–299.

Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics*, 44, 193–215.

Jimenez–Contreras, E., Lopez–Cozar, E.D., Ruiz–Perez, R., Fernandez, V.M., (2002). Impact–factor rewards affect Spanish research. *Nature*, 417, 898–898.

Kindermann G., (1999). Hat Deutsch noch Zukunft als Wissenschaftssprache? *Geburtshilfe und Frauenheilkunde*, 59, 188–190.

KNAW Commission, chaired by Professor van Bemmel *"Kwaliteit verplicht",* report. Amsterdam, 2000.

van Leeuwen T.N., Moed, H.F., Tijssen, R.J.W., Visser, M.S., van Raan, A.F.J., (2000). First evidence of serious language–bias in the use of citation analysis for the evaluation of National Science Systems. *Research Evaluation*, 9, 121–122.

van Leeuwen T.N., Moed, H.F., Tijssen, R.J.W., Visser, M.S., van Raan, A.F.J., (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics*, 51, 335–246.

van Leeuwen, T.N., Visser, M.S., Moed, H.F., Nederhof, A.J., van Raan, A.F.J., (2003). The Holy Grail of Science Policy: exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics*, 57, 257–280.

Luwel, M., Moed, H.F., Nederhof, A.J., De Samblanx, V., Verbrugghen, K., van der Wurff, L.J., (1999). *Towards indicators of research performance in the social sciences and humanities.* An exploratory study in the fields of Law and Linguistics at Flemish Universities. Report of the Flemish Inter-University Council (V.L.I.R.), Brussels, Belgium / Centre for Science and Technology Studies (CWTS), Leiden University, the Netherlands / Ministry of the Flemish Community, Brussels, Belgium. V.L.I.R. Brussels, Belgium.

Moed H.F., *Thesis*, Leiden University, 1989.

Moed, H.F., De Bruin, R.E., van Leeuwen, T.N., (1995). New bibliometric tools for the assessment of national research performance: database description, overview of indicators and first applications. *Scientometrics*, 33, 381–422.

Moed H.F., (2002). The impact factors debate: the ISI's uses and limits. *Nature*, 415, 731–732.

Narin, F., (1976). Evaluative bibliometrics. *The use of publications and citation analysis in the evaluation of scientific activity*. Computer Horizons, Inc., Cherry Hill, New Jersey, USA.

Nederhof, A.J. (1988) *The validity and reliability of evaluation of scholarly performance*. In A.F.J. van Raan (Ed.), Handbook of Quantitative Studies (pp. 193–228). Amsterdam: North-Holland.
NOWT, *Science and Technology Indicators report 1994* (in Dutch).
NOWT, *Science and Technology Indicators report 1996* (in Dutch).
NOWT, *Science and Technology Indicators report 1998* (in Dutch).
NOWT, *Science and Technology Indicators report 2000* (in Dutch).
NOWT, *Science and Technology Indicators report 2003* (in Dutch).
OST, *Science & Technologie Indicateurs – Edition 1992*, Economica, Paris, 286 pp.
OST, *Science & Technologie Indicateurs – Edition 1994*, Economica, Paris, 425 pp.
OST, *Science & Technologie Indicateurs – Edition 1996*, Economica, Paris, 473 pp.
OST, *Science & Technologie Indicateurs – Edition 1998*, Economica, Paris, 551 pp.
OST, *Science & Technologie Indicateurs – Edition 2000*, Economica, Paris, 512 pp.
OST, *Science & Technologie Indicateurs – Edition 2002* ,Economica, Paris, 467 pp.
PREST (2000). *Impact of the Research Assessment Exercise and the Future Quality Assurance in the light of changes in the research landscape*. Report to the Higher Education Funding Council for England (HEFCE), by PREST, University of Manchester, United Kingdom.
RAE 2001. *Guide to the 2001 Research Assessment Exercise.*
Van Raan, A.F.J. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight studies. *Scientometrics*, 36, 397–420.
Rempen, A., (1998). Leserbrief zur Haller, U., H. Hepp and E. Reinhold Tötet der 'Impact factor' die deutsche Sprache*? Gynakologisch–geburtshilfliche Rundschau*, 38, 54–54.
Saiz–Salinas, J.I., (1996). Failed professor. *Nature*, 381, 186.
Tijssen, R.J.W., van Leeuwen T.N., van Raan, A.F.J. (2002). *Mapping the scientific performance of German medical research. An international comparative bibliometric study*. Report to the German Federal Ministry of Education and Research (BMBF).
VSNU, *Standard Evaluation Protocol (1998) for Public Research organisations*, Utrecht, 1998.
VSNU, NWO, and KNAW, *Standard Evaluation Protocol (2003 – 2009) for Public Research organisations*, January 2003.
Warner, J. (2000a). A critical review of the application of citation studies to the Research Assessment Exercises, *Journal of Information Science*, 26, 453–460.
Warner, J. (2000b). Research assessment and Citation Analysis, *The Scientist*, 14, October 30.

Chapter 17

WHAT HAPPENS WHEN FUNDING IS LINKED TO PUBLICATION COUNTS?

Linda Butler

Research Evaluation and Policy Project, Research School of Social Sciences, The Australian National University, Australia
E-mail: linda.butler@anu.edu.au

Abstract: Many countries are placing a greater emphasis on public accountability for government research funding and are starting to use quantitative performance indicators for the distribution of funds. In Australian universities the use of quantitative formulas to allocate the research component of university block grants to institutions has been in place for a decade, and thus the system provides fertile ground for using bibliometrics to examine the effects of such policies on academic output. An analysis of Australian data from the Institute for Scientific Information's major citation indexes clearly demonstrates the academic response to the linking of funds, at least in part, to productivity measures undifferentiated by any measure of quality — publication numbers jumped dramatically, with the highest percentage increase in the lower impact journals. The trends were apparent across all fields of research in the university sector, but were not present in other sectors active in research (such as hospitals or government research agencies). The trends were not, however, uniform across all institutions.

1. INTRODUCTION

In most OECD countries increasing emphasis is being placed on greater public accountability, with a need to demonstrate the effectiveness and efficiency of government supported research. A workshop held by the OECD in 1997 characterised the recent evaluation of basic research as "a rapid growth industry"(OECD, 1997).

This new demand for research evaluation cannot be fully serviced by the finite capacity of traditional peer review. Researchers, particularly the more

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 389-405.

senior ones, have many calls on their expertise, such as reviewing journal articles, assessing grant applications, sitting on selection and promotion committees, being co-opted to national or institutional review bodies. They can only devote a limited proportion of their time to such activities before their own research begins to suffer. Partly as a consequence of the pressures on peer review, there has been an increased use of quantitative performance indicators as an alternative method for evaluating research performance, which has the added advantage of being more cost efficient. There is also an increasing trend to link such measures directly to the distribution of research funds.

For Australian universities the allocation of funds earmarked for research is based on a formula encapsulating a number of performance measures (graduate student numbers or completion rates, research income, and publications). Spanish scientists are directly rewarded with a salary supplement for increasing their output in the major English language international journals (Jiménez–Contreras, Anegón and López–Cózar, 2003). In Finland part of the funding for university hospitals rests on publication points, weighted according to the impact factor of the journals carrying the work (Adam, 2002). While in the British Research Assessment Exercise the link between research rankings and performance measures, and hence funding, is less direct, they nevertheless play an important role in the deliberations of the review panels.

The link between research funding and quantitative performance measures has now been in place in Australian universities for a decade, and thus provides fertile ground for using bibliometric data to examine the effects of this policy on academic output. Since performance measures relating to publications are limited to aggregate productivity counts, the expectation would be that Australian university publication output would increase significantly in response to the signals embodied in the funding formula. As there is no attempt to weight for the quality of either the output itself, or the publication in which it appears, there would also be an expectation that any increased journal output is likely to be concentrated in lower ranked journals where it may be easier to place additional articles. Both these anticipated outcomes are clearly visible in the data for Australian universities in a number of major journal citation indexes.

2. POLICY BACKGROUND

The Australian government has a dual system for funding research in universities. A significant amount of money is distributed by the two research councils, the National Health and Medical Research Council and the Australian Research Council, via a peer reviewed assessment system. Both agencies distribute the bulk of their funding support in the form of project grants, which can vary in length from one to five years, with three years being the most common duration. Secondly, a proportion of the block operating grant to universities (of the order of 5%) is earmarked for research and research training, and since the beginning of the 1990s this has been distributed via a formula. The formula aimed at taking account of a broad range of measures of research performance when making allocations to universities. Initially this formula was based only on external earnings, but subsequently student and publication components were added.

Australian universities began supplying details of their research output to the Department of Education, Science, and Technology (DEST[1]) and its predecessors in 1993, initially through the Australian Vice Chancellors Committee (AVCC), and more recently directly to the department. The research funding formula was expanded in 1995 to include output measures — publication counts and higher degree loads and completions — and was also used in the allocation of postgraduate awards. The components of the formulas, the funding schemes they were applied to, and the weighting given to each element, are shown for a sample of years in Table 17.1.

From 2001, as a result of a review of higher education research, the amount of funds allocated on the basis of formulas has nearly trebled, and now accounts for more than half the funding specifically targeted to research and research training through the education portfolio (DEST, 2002a). The Small Grants scheme, not previously funded by this method, was rolled in with the Research Quantum (RQ) and became the Institutional Grants Scheme. Postgraduate awards continued to be funded under this arrangement and, in addition, a new Research Training Scheme was introduced which more than doubled the funds distributed via formulas. None of the more recent changes represented 'new' money, merely a change in the method by which some of the funds were distributed, and a greater reliance on formula driven schemes.

Australia's approach in this area of higher education policy is not common. A recent survey of 14 countries by Geuna and Martin only identified two that used ex-post quantitative evaluation for allocating core research funds, Finland and Australia (Geuna and Martin, 2003) Unlike Australia's mechanistic system of quantitative measures, Finland employs a

[1] The Australian Government department which encompasses the education portfolio has had several name changes in the period referred to in this paper — the Department of ... Employment, Education, Training, and Youth Affairs; Employment, Education, and Training; and Education, Science, and Training — but I will use the acronym for the department in its current form (DEST) throughout this chapter

series of agreed indicators focusing on the quality and impact of teaching and research. The Australian experience is not mirrored in other countries, and may well be part of the explanation for the publication trends seen in Table 17.1.

Table 17.1. Formulas that distribute research funds to Australian universities through block grants

	Weight given to each element (percent)				
Funding Scheme	Total funds ($mil)	Publications	Higher degree load	Higher degree completions	Research income
1996					
Research Quantum	218.6	12.50		5	82.50
Postgraduate awards (2 schemes)	91.7	5.26	40	20	34.74
2000					
Research Quantum	223.0	10.00		10	80.00
Postgraduate awards (2 schemes)	96.2	4.44	40	20	35.56
2002					
Institutional Grants Scheme	271.3	10	30		60
Postgraduate awards (2 schemes)	102.0	10		50	40
Research Training Scheme	515.6	10		50	40

Source: Australian Vice Chancellors' Committee (AVCC), 2002.

3. THE REWARDS FOR PUBLISHING

Determining the 'value' of a publication unit to a university is a simple calculation and it was not long before figures became commonly referred to in the sector. Taking the data given in Table 17.1, together with the publication counts on which the distribution of funds was based, Table 17.2 details the calculations for the three sample years. The distribution of funding for the publications element was based on data for the most recently available two years.

Table 17.2 demonstrates the effect that adjustments to the coverage of publications in the collections, and/or the amount of funding distributed in this way, can have on calculations of the unit value. For example, the 1996 distribution was based on 1993 and 1994 publications. The 1993 data covered 8 publication types; the 1994 data covered 22. After a sample audit of the universities' lists of 1994 publications, the number of categories covered was reduced to just four for subsequent collections: books, book chapters, refereed journal articles, and refereed conference papers[2]. As a result the number of publication units in subsequent collections dropped significantly, with a consequential increase in the value of each unit. This occurred despite a reduction in the weight given to publications in the formula from 12.5% to 10%, and a reduction in the amount of funds distributed on this basis.

Table 17. 2. Value of a publication unit: 1996, 2000 and 2002

Funding year	*Funds tied to publication counts (AUD$million)*	*Publication counts**	*Value per publication unit*
1996	32.1	42,259	$761
2000	26.6	24,390	$1,089
2002	88.9	26,877	$3,307

Source: Department of Employment Education and Training, 1996; Department of Education, Training and Youth Affairs (DETYA), 2000; Department of Education, Science and Training (DEST), 2002b.
* Weighted by type of publication

From 2001 the funds distributed via the formulas were increased significantly, leading to a three-fold increase in the value of a publication unit. Every refereed journal article is now 'worth' over AUD$3,000 to a university, and a book is now 'worth' AUD$15,000.

4. IDENTIFYING THE EFFECTS OF INTRODUCING FUNDING FORMULAS

As the categories covered by the Australian collection have been refined and reduced in number, the importance of journal publications indexed by the *Institute for Scientific Information* (ISI) has increased. The collection is externally audited, and universities must prove, among other things, that the journals carrying the articles they are claiming are peer reviewed. A journal that is indexed by ISI is accepted as peer reviewed without question, but universities must prove that any other journal meets the definition. Publishing in ISI-indexed journals is obviously the easiest course of action to take. The data contained in ISI's three main indices, the *Science Citation*

[2] In recent collections books receive a weighting of five in the calculations, while the other three categories are all given the base weighting of one

Index (SCI), the Social Sciences Citation Index (SSCI), and the Arts and Humanities Citation Index (A&HCI), therefore provide fertile ground for examining the impact that introducing the funding formulas had on Australian university output.

The Research Evaluation and Policy Project (REPP) maintains a database which contains all Australian publications in these ISI indices. Considerable effort is expended in standardising the addresses listed for each publication, thus enabling accurate analysis to be undertaken at the sectoral (university, hospital, government, etc), institutional, and even lower levels of aggregation, such as faculties and departments.

4.1 The University Sector in Aggregate

An analysis of Australia's presence in the SCI was the first step taken to investigate whether it was possible to demonstrate the apparent effect of the introduction of the funding formulas in the 1990s. In the analysis SCI journals were allocated to quartiles based on the average citation per publication rates of the publications they carried. Mean journal citation rates were calculated for each five year window from 1981–85 through until 1996–2000. For both publication counts and citation totals, the calculation was limited to publications classified by ISI as articles, notes, reviews and proceedings papers, and to the specified five year period. As a separate calculation was made for each period, journals were free to move between quartiles over time.

Australian universities' presence in these four quartiles was then tracked over the full twenty year period. Their share of total publications in each of the four quartiles is shown in Figure 17.1.

The response of the academic community appears very clear, and in line with expectations. Until the period 1989–93 there had been virtually no movement in the institutions' presence in the SCI journal set, with the exception of an increase in the third quartile. Since that period university output has jumped dramatically, particularly in journals allocated to the bottom two quartiles. The sector's share of publications in journals allocated to the top two quartiles increased by 28% and 15% respectively; their share of publications in the third quartile rose at double those rates, i.e., by 55%; and in journals allocated to the bottom quartile their share doubled.

With no attempt made to differentiate between the quality, visibility or impact of the different journals when funding is allocated, there is little incentive to strive for publication in a prestigious journal. Whether a publication reports ground breaking research or is a more pedestrian piece; whether it appears in a highly visible journal such as *Nature* or a lower impact outlet, the rewards are identical.

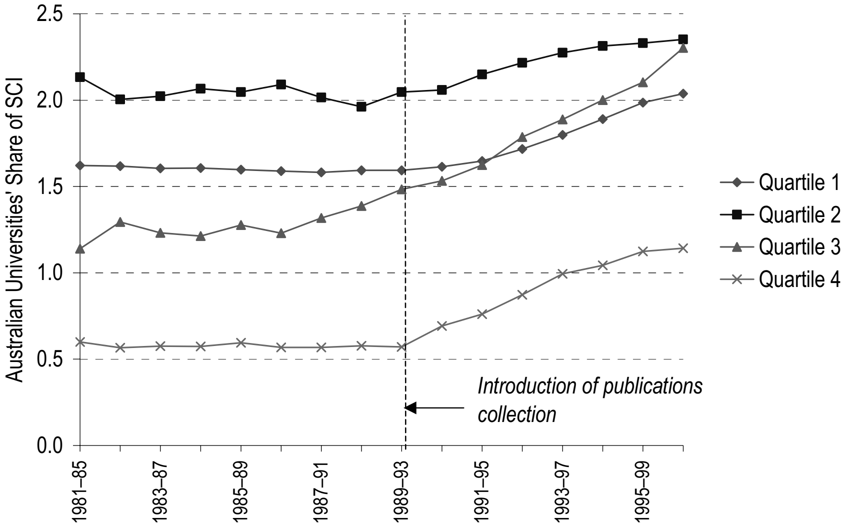


Figure 17 1. Australian universities' share of publication in the SCI, by journal impact quartile: five year windows, 1981–1985 to 1996–2000.

The trends shown in Figure 17.1 are not proof in themselves of a direct link between funding formulas and increased productivity. However, they did occur at a time when funds to the sector are extremely tight. A detailed analysis was undertaken when these trends first became apparent to determine whether the increased output could be explained by either the entry of new institutions into the sector, or an increased number of researchers (Butler, 2001a) Results showed that while the new institutions had increased the sector's research capacity, they accounted for less than one third of the expanded output — the bulk came from the older, established universities. Nor were increased staff numbers the explanation. They had risen in the period after the introduction of the publications collection, but the increase was no greater than it had been prior to this time.

To be more confident that the trends are a result of the introduction of funding formulas it is necessary to examine the data in more detail in order to determine whether the following three scenarios also exist:

1. The trends are specific to the university sector. No other Australian research sector is faced with the same funding drivers, so the trends for other sectors should not mirror that for universities.

2. The trends are present in all fields of research. The formulas are applied across the university sector, so all fields of research, including those less traditionally reliant on journal outlets for their research, should exhibit similar trends.
3. Another university system faced with similar incentives, exhibits similar trends. The Spanish research system is subject to funding drivers based on journal output, and the effect of this should also be apparent in ISI data.

The results of these analyses are given in the following sections.

4.1.1 Comparison of sectors

The three largest sites of research in Australia outside the universities are the Commonwealth Scientific and Industrial Research Organisation (CSIRO), the hospitals, and government research agencies. None of the institutions comprising these sectors are subject to funding formulas of the type present in the university sector, although all have strong collaborative links with it. Figure 17.2 shows the trend in publication output for these three sectors using an identical analysis to that applied to the university sector.

It is quite clear that the 1989–93 period does not mark a turning point in trends for any of these sectors. CSIRO, with an increasing emphasis on seeking external funds for a significant share of its operating costs, has seen its overall share decline (although actual publication numbers have remained steady). The hospital sector's share of output in the top quartile has been increasing steadily across the whole period, while its presence in the journals allocated to the bottom quartile has increased but remains very low. There are considerable fluctuations in its share of the other two quartiles, and the mirror image in movement between these quartiles suggests some journal movement between the two sets. The government sector's share of output in the top and bottom quartiles has remained relatively steady across the twenty year period covered by our data. As with the hospital sector, their presence in quartiles 2 and 3 is more volatile, and presents a mirror image in movement.

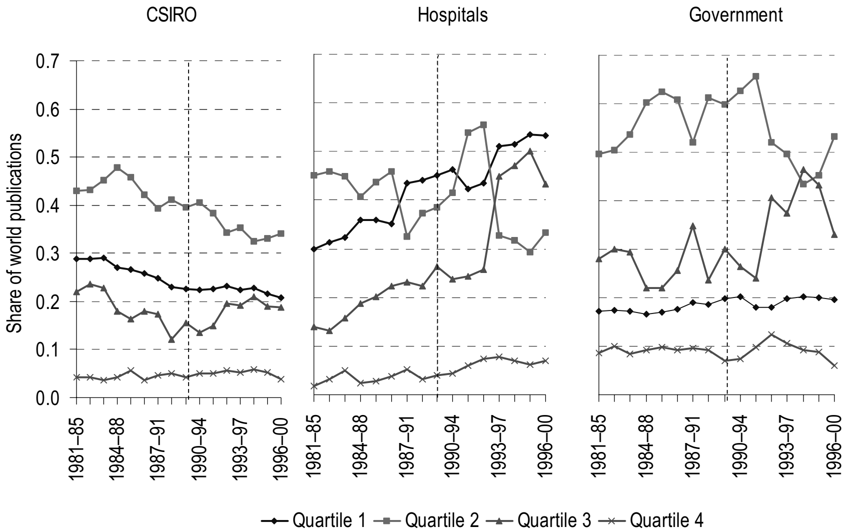


Figure 17. 2. Share of publications in the SCI by other Australian sectors, by journal impact quartile: 1981–1985 to 1996–2000.

4.1.2 Comparison of fields

To disaggregate the trends and examine what was occurring in different fields of research, the methodology used for the SCI as a whole was applied to subsets of journals. For this analysis ISI subject category journal sets were used, and translated into the Australian Research fields, Courses and Disciplines classification scheme. Within each field journals were allocated to quartiles on the basis of the five year average citation impact of the publications they carried. As expected, the average citation per publication (cpp) threshold varied considerably between fields. For example, to be in the top quartile in chemistry in the period 1996–2000, a journal needed a cpp rate of 3.61, while a mathematics journal required only 1.86.

Table 17.3 shows the increase in Australian universities share of world publications by field in two periods of equal length: the increase between 1981–85 and 1988–92; and the increase between 1989–93 and 1996–2000, the period after the introduction of the publications collection.

Table 17 3. Percentage increase in publication output by field — two periods

	% Change: 81–85 to 88–92				*% Change: 89–93 to 96–00*			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
All sciences	–2	–8	22	–4	28	15	55	100
Mathematical sciences	–13	–16	–3	14	1	43	34	77
Physical sciences	–8	–25	137	–32	42	63	18	85
Chemical sciences	–20	–8	47	13	24	–17	124	137
Earth sciences	7	15	38	19	4	28	31	88
Biological sciences	–7	–5	41	–17	18	25	27	74
Engineering and technology	–10	0	16	–1	37	42	75	117
Agric, vet, environ	–16	14	78	–21	14	48	52	144
Medical and health sciences	0	9	0	29	22	18	84	82
Social sciences	4	–19	56	–25	13	63	28	65

Most fields of research demonstrate relatively stable publication shares between 1981–85 and 1988–92, with movements contained within 25%. The exception is increases in the third quartile — in line with overall trends.

These data show, with the one exception of a decrease in university publications in the second quartile in chemistry, that universities have significantly increased their output in all fields and in all quartiles in the second period studied (1989–93 to 1996–00). In the medical and health sciences the increase in share of the bottom two quartiles is at a similar level; in all other fields the largest increase is in the bottom quartile, usually by a significant margin. As with trends in the preceding period, in this later time frame the fields exhibit trends similar to the aggregate ones, although inevitably there is some variation. In most cases quartile 3 accounted for the second largest increase, with those in the top two quartiles much more modest.

The increase in output in the physical sciences is more evenly spread across the four quartiles. Notably, universities increased their share of the highest impact journals by 40%, a greater margin than for any other field. The two possibilities which immediately suggested themselves as an explanation for this trend — the influence of astronomy in which Australia is particularly strong, and the movement of major Australian journals in the field between quartiles — were found to have no impact on the trends.

4.1.3 The Spanish experience

Since 1989 a research incentive system has existed in Spain, administered by the National Commission for the Evaluation of Research Activity (CNEAI). Researchers were rewarded with salary bonuses for publishing in prestigious journals, principally articles appearing in a relatively high position (approximately the top one third) in ISI's Journal

Citation Report lists by subject category. Unlike the Australian system, the focus is clearly on the individual rather than the institution. But the message is clear — it is increased productivity that is important. A recent study has clearly demonstrated the effect of this policy on Spanish publication output in the ISI–indexed journals (Jiménez–Contreras, Anegón and López–Cózar, 2003).

Their work demonstrates clearly that Spanish researchers have also responded to funding stimuli by increasing their output well above the long-term trend line for Spanish publications in the ISI indices. However, in the Spanish case CNEAI achieved its stated aims, which were to increase productivity and the internationalisation of Spanish research. In contrast, the Australian funding formulas were designed to reward quality, but in fact reward quantity.

4.1.4 Interpretation of trends

The similar trends found in university output in all fields of research, the lack of similar trends in other research sectors, and the Spanish experience, all support the hypothesis that the increased university output in Australia, and the pattern of its distribution across impact quartiles, is a direct result of the introduction of the DEST funding formulas.

There are differing interpretations which can be placed on these trends. In discussions which followed the release of the data, there were those who argued it was 'good news' — that the large jump in output in low impact journals was of little concern because the Australian presence in high impact journals had also increased. While this may be true, there is an overriding objection to the use of undifferentiated publication counts in this instance, and that is one of intent. The formulas, and in particular the publications component, were conceived as a means of distributing research funds on the basis of the quality of research in Australian universities. Publication counts are not measures of quality.

4.2 Institutional Analysis

While the trends in Australian publication output were similar across the different fields of research, it is perhaps not surprising that trends in individual institutions are not as uniform. This is largely because of the disparate signals which individuals within these institutions are receiving from a variety of sources, their judgment on which carry the most weight, and their subsequent reaction to these signals. Researchers face one set of performance measures when applying for grants; another when seeking promotion; yet another when applying for a job at a new institution; a series of community standards set by the peers in their own discipline — all in addition to any sector–wide signals which their institution may be receiving and passing on down through faculties and departments. Some of the signals received will inevitably be contradictory.

Table 17.4 shows publication trends for individual institutions calculated in the same manner as for fields. To provide some indication of the nature of each institution the universities have been classified by type and by the size of their output for the two periods. Australian universities are often classified into four categories:

– 'Go8' (Group of Eight) universities are a self–selected group with a strong research focus and a wide coverage of disciplines. Most are among the oldest of the nation's universities, the first institutions to be established in the major State capital cities. The exceptions are New South Wales, Monash, and the Australian National University, although all three have been established over 50 years;
– 'pre-1988' universities are more recent, but were in existence prior to the major higher education reforms of 1988 which saw the abolition of Institutes of Technology or Colleges of Advanced Education as distinct types of tertiary institutions;
– 'ex-IT' universities are those which, prior to the 1988 reforms, were solely undergraduate institutes of technology. A few of the larger, older establishments had already been granted university status just prior to the major reorganisation of the sector; and
– 'ex-CAE' universities are those which, prior to the 1988 reforms, existed primarily as small, undergraduate institutions focusing on the professions, such as teaching and nursing, with little research capacity.

Table 17.4 has been limited to those institutions with at least a modest publication profile in the 1980s — those with less than 100 publications in the five year period 1988–1992 were excluded.

The institutions with the greatest overall increase in publication output are the 'ex-CAEs' and the 'ex-ITs'. For both groups this is to be expected, because their capacity to undertake research, and the number of staff qualified and experienced to do so, increased significantly after the change in status of their institutions.

Only four institutions showed a greater growth in publication output in the first period (1981–85 to 1988–92) than in the second period (1989–93 to 1996–00). Two were 'ex-CAEs' which started from a low publication base — University of Western Sydney and Northern Territory University. The other two institutions were 'pre-1988' universities — Deakin University and University of New England. Deakin University's publication trends are unique among Australian universities with more growth in the earlier period, and the highest increase in the second period to be found in the top quartile.

All other institutions in the analysis showed a significantly greater growth in publication output in the second period. In fifteen instances the highest growth rate was in the bottom quartile, while in another four cases the highest growth was recorded in quartile three. In the remaining five cases four recorded their highest growth in quartile two and just one institution, James Cook University, recorded its strongest growth in the top quartile.

Table 17. 4: Publication output trends for Australian universities — two periods

	Type	*No. pubs*	*% change: 81–85 to 88–92*				*No. pubs*	*% change: 89–93 to 96–00*			
University		*88–92*	*Q1*	*Q2*	*Q3*	*Q4*	*96–00*	*Q1*	*Q2*	*Q3*	*Q4*
All Universities		37,721	–2	–8	22	–4	60,014	28	15	55	100
U Sydney	Go8	10,620	20	41	36	13	17,628	51	54	71	93
UAdelaide	Go8	6,048	1	83	12	3	8,350	23	10	64	206
Australian Natl U	Go8	5,595	–8	49	6	–3	7,536	35	24	42	71
U Queensland	Go8	3,987	33	42	20	–6	7,514	78	55	103	82
U Melbourne	Go8	5,170	22	37	28	–4	7,490	28	37	74	104
New S Wales	Go8	4,270	28	71	10	15	6,628	35	45	60	120
Monash	Go8	3,438	5	57	–6	3	5,386	49	33	67	103
U W Australia	Go8	3,054	1	70	74	26	5,052	47	60	48	112
Queensland U Tec	ex–IT	668	4	243	126	79	2,554	202	162	150	453
La Trobe U	pre1988	1,643	0	81	18	3	2,235	15	43	33	63
Flinders U	pre1988	1,930	13	17	36	–2	2,119	–3	6	77	76
U Tasmania	pre1988	1,143	16	71	14	23	2,021	60	66	69	128
U Newcastle	pre1988	1,222	12	65	7	0	1,891	17	47	89	158
Macquarie U	pre1988	1,138	35	87	18	–21	1,700	41	88	30	30
U Wollongong	pre1988	697	40	100	56	50	1,537	73	168	86	70
James Cook U	pre1988	680	1	95	7	144	1,451	115	51	95	94
Griffith U	pre1988	756	–2	116	65	100	1,250	35	66	122	76
U New England	pre1988	1,059	2	46	13	3	1,115	–5	–2	–10	17
Curtin U	ex-IT	409	36	52	50	120	1,113	56	156	154	241
Murdoch U	pre1988	805	–23	61	93	97	1,064	8	9	114	23
Deakin U	pre1988	491	44	115	142	97	880	80	75	56	68
RMIT	ex-IT	261	10	144	10	53	821	91	212	248	248
U Tec Sydney	ex-IT	250	25	116	4	40	772	97	174	168	172
U South Australia	ex-IT	170	53	15	59	–8	695	113	436	251	174
U Western Sydney	ex-CAE	155	617	180	750	375	680	275	289	197	305
Charles Sturt U	ex-CAE	113	20	157	88	29	364	226	248	173	400
Nth Territory U	ex-CAE	122	1150	2400	200	283	279	133	120	117	20
U Canberra	ex-CAE	128	21	293	58	–55	272	235	70	49	242

The universities with the most even growth across quartiles subsequent to the introduction of the funding formulas were University of Sydney, Australian National University, University of Queensland, La Trobe University and Deakin University, all with a standard deviation of less than 20.

5. DISCUSSION

Problems with the composite index, and in particular with the publications component, were raised soon after its introduction (Anderson, Johnson and Milligan, 1996). Most of the discussion concentrated on the Research Quantum (RQ) as it was the largest scheme. These concerns were taken on board in a ministerial discussion paper on higher education research and research training, issued in June 1999:

> "The publications component of the Composite Index has been subject to a range of criticisms since its implementation in 1995. These concern the reliability of the information provided by institutions, the costs of data collection and the incentives created by the inclusion of a publications component in the index. It seems likely that the publications component of the Composite Index has stimulated an increased volume of publication at the expense of quality … on these grounds, the Government proposes ... to drop the publications measure in any future indices used to allocate block research funds" (Kemp, 1999a).

Not all universities were keen to see the removal of the publications element. The notional proportion of the RQ to be distributed via the publications component was 10% in 1999. However, over half the universities, particularly smaller institutions, received more than 10% of their RQ allocation through publications. For one university the proportion was above 40%; for another five it was more than 20%. It was predominantly the research intensive older universities that were at, or even under, the 10% benchmark (DEST, 1999).

It was therefore hardly surprising that in its response to the discussion paper, the AVCC, representing all 36 institutions which received funds via the RQ, argued for the retention of the publications component:

> "… of the quality measures that might be utilised, 'publications' is the only measure able to fulfil all the requirements … for a driver of sector–wide funding" (AVCC, 1999).

The government was swayed by the submission of the AVCC and others, and in its final policy statement all talk of removing the publications component had disappeared (Kemp, 1999b).

Concerns also surfaced about the direction in which the publications component of funding formulas currently in place in the higher education sector was driving universities, when data produced by the *Institute for Scientific Information* (ISI) confirmed the marked increase in Australian output in the journal literature but pointed to a significant decline in citation impact relative to many OECD countries (Butler, 2001b).

The concerns raised back in 1999 about the use of an undifferentiated publication count are re-surfacing in the context of the latest review of the Australian higher education system. A number of submissions to the government review established to evaluate the Knowledge and Innovation reforms have suggested the removal or modification of the publications component (DEST, 2002c). Two questions stakeholders were specifically asked to address related to the publications collection:

> "Should the research publications element be removed from the formulae? Should the research publications element of the formulae include quality measures"(DEST, 2002c).

In their submissions the majority of institutions remain committed to the continuation of the collection. A number would like to see the introduction of quality measures, although generally this approach has been rejected because few have any knowledge of possible performance measures that could be used to approximate the notion of quality. Most appear to assume it means weighting publication counts by ISI's journal impact factor, or using aggregate citation counts, and have no knowledge of the more complex and sophisticated bibliometric methods that have been developed in recent years.

The University of Central Queensland highlights another problem with the collection in its existing form:

> "The resources used in collecting, submitting and verifying publications by institution exceed the income received for publications at Central Queensland University".

It is clearly apparent that before any alternative could be adopted there needs to be detailed assessment of possible measures. Several questions need to be examined. Do the proposed indicators come close to measuring the aspect of the research endeavour the government is targeting? Is the measure suitable for the level of aggregation being assessed? Is the measure applicable to all fields of research? Is the necessary data readily available and independently verifiable? Is it more effective to combine a suite of indicators, perhaps varying by field, rather than relying on a single measure? Only when these and other questions have been answered, and the effects of their introduction demonstrated, would there be any chance of gaining broad acceptance for the replacement of existing measures.

ACKNOWLEDGEMENTS

Certain data included herein are derived from the *Australian National Citation Report* prepared by the *Institute for Scientific Information®*, Inc. (ISI®), Philadelphia, Pennsylvania, USA: ©Copyright Institute for Scientific Information® 2000. All rights reserved. This chapter updates and extends research first reported in *Research Evaluation* in 2003 (Butler, 2003).

REFERENCES

OECD. (1997). *The evaluation of scientific research: selected experiences*. Paris: OCDE/GD(97) 194.

Jiménez–Contreras, E., Anegón, F.M., López–Cózar, E.D. (2003). The evolution of research activity in Spain: the impact of the National Commission for Evaluation of Research Activity (CNEIA). *Research Policy*, 32 (1), 123–142.

Adam, D. (2002). The counting house. Nature, 415, 726–729.

Australian Vice Chancellors' Committee (AVCC). (2002). Time series data relating to DEST higher education research data collection (http://www.avcc.edu.au/ australias_unis/ statistics/research/index.htm).

Department of Education, Science and Training (DEST) (2002a): (http://www.detya.gov.au/ highered/research/index.htm#funding.

Guena, A., Martin, B.R. (2003). University Research Evaluation and funding: an international comparison. Minerva, 41 (4), 277–304. http://www.sussex.ac.uk/spru/publications/ imprint/sewps/sewp71.html.

Department of Employment Education and Training. (1996). *Higher education report for the 1996 to 1998 triennium*. Canberra.

Department of Education, Training and Youth Affairs (DETYA). (2000). *Higher education report for the 2000 to 2002 triennium.* Canberra.

Department of Education, Science and Training (DEST).(2002b). *Higher education report for the 2002 to 2004 triennium*. Canberra.

Butler, L. (2001a). *Monitoring Australia's scientific research*. Canberra: Australian Academy of Science, 20.

Anderson, D., Johnson, R., Milligan, B. (1996). *Performance–based funding of universities*. Canberra: Commissioned Report No.51, National Board of Employment Education and Training.

Kemp, D. (1999a). *New knowledge, new opportunities*. Canberra: Department of Education, Training and Youth Affairs.

Department of Education, Science and Training (DEST) (1999). *Characteristics and performance indicators of higher education institutions 1999*, http://www.detya.gov.au/archive/highered/statistics/characteristics/30_researchquantum.htm.

AVCC. (1999). *Discussion paper on higher education research and research training*. Canberra: AVCC.

Kemp, D. (1999b). *Knowledge and Innovation*. Canberra: DETYA.

Butler, L. (2001b). *Monitoring Australia's Scientific Research*. Canberra: Australian Academy of Science, p.11.

Department of Education, Science and Training (DEST) (2002c). *Setting firm foundations: financing Australian higher education*. Canberra: DEST.

DEST. *Evaluation of knowledge and innovation reforms: issues paper*. http://www.detya.gov.au/highered/ki_reforms/issues_paper.rtf.

Butler, L. (2003). Modifying publication practices in response to funding formulas. *Research Evaluation*, 12 (1), 39–46.

Chapter 18

INTERNATIONALISATION IN SCIENCE IN THE PRISM OF BIBLIOMETRIC INDICATORS

Journals, Collaboration, and Geographic Distribution

Michel Zitt[1,2] and Elise Bassecoulard[1]

[1]*INRA, Lereco, Institut National de la Recherche Agronomique, Nantes, France*
[2] *OST, Observatoire des Sciences et des techniques, Paris, France. E-mail: zitt@obs-ost.fr*

Abstract: Powerful engines tend to support internationalisation: self-organisation of scientific communities regardless of national borders; international and supra-national top down programmes; side effects of economic globalisation; all these trends being boosted latterly by the ICT/Internet revolution. However, internationalisation meets several obstacles: resistance of the national structure in most aspects of innovation systems; proximity effects anchored in infra-structural factors; inertia of personal and institutional networks. Internationalisation of competition and cooperation does not necessarily imply fewer discrepancies in national performances. Bibliometric studies of scientific journals profiles, collaborative and other scientific networks, spatial distribution of scientific activity, tend to validate a real but slow process of the fading of borders. In the last decade advances appear more in globalisation of scientific communication and increase of aggregate collaboration figures than in the geographic distribution of knowledge sources, the reshaping of co-operation networks and the modification of interdisciplinary balances in connection with new growth regimes of science.

1. INTRODUCTION

Internationalisation in science is often taken for granted. Powerful engines tend to support internationalisation: self-organisation of scientific communities regardless of national borders, international and supra-national top down programmes, side effects of economic globalisation, all these trends being boosted latterly by Information and Communication Technology (ICT) progress and Internet revolution. However, internationalisation meets several obstacles: resistance of the national factor

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 407-436.

in most aspects of innovation systems, proximity effects anchored in infrastructural factors, inertia of personal and institutional networks. This context is recalled in section 2. Results of these antagonistic mechanisms can be empirically studied. As far as outputs of scientific activities are concerned, bibliometric measures help to assess the degree of achievement of various forms of internationalisation, not necessarily convergent: reduction of barriers to competition; international co-operation and coordination; reduction of inequalities in scientific output. Section 3, devoted to the internationalisation of media, especially the scientific journals, exemplifies the first form. Section 4 addresses the internationalisation of interdependence networks, with a focus on international collaboration. Section 5 addresses the process of homogenisation and convergence of scientific production. The final section is devoted to discussion and conclusions.

2. BACKGROUND

A Republic of Science unconstrained by political borders has been a dream of many scientists and various steps of progress toward a self-organisation of communities beyond national and cultural differences have been celebrated by observers of research as a social object, among many others Merton (1973) and Price (1963).

2.1 Engines of Internationalisation

History of science teaches that scientists consider it natural and profitable to freely communicate and collaborate, and professionalisation of science in the XIXth and XXth centuries has fostered this trend (DeBeaver and Rosen, 1978). This self-organisation is the first engine of science internationalisation. The creation of the Nobel Prize in 1901 was a symbol of internationalism in the Republic of Science (Crawford, 1992). The history of physics in the early 20th century, for example, is tightly linked to the development of international meetings and exchanges, despite nationalist pressures' interference. Border–free competition and co-operation are at the heart of the self-organisation of science.

The second engine of internationalisation lies in the top down processes, which gained full power after WWII. Multinational programmes associate clubs of countries with reasonably converging political objectives, either on an occasional or permanent basis. Top down processes and self-organisation interact in many ways in large scale programs: cost sharing of large facilities[1] (physics/ astrophysics), co-ordination of large programs (the genome). Supra-national entities, first of all the EU, became an important source of coordination and funding of programs aiming at convergence and integration of member countries, with a heavy Science–Technology–Innovation folder. Framework programmes, incentives to networking and mobility, efforts to harmonise higher education systems, and currently the European Research Area initiative are expected to enhance the cohesion and competitiveness of Europe.

A third engine has gained force in the last decades, namely the general movement of financial and economic globalisation. It has been celebrated as a strong mechanism of diffusion of knowledge, in particular through multinational firms. R&D services' implementation and their articulation with local research are often viewed as an important internationalisation engine. Academic research is enrolled through tighter linkages with technology and markets. An echo of the increasing pressures on Mertonian model is found in the 'new economics of science' (Dasgupta and David, 1994, David and Foray, 1995, Stephan, 1996; for a contrasting view see Callon, 1994).

Bringing drastic cuts in communication costs, the new Information and Communication Technology (ICT) and Internet revolution has boosted non-physical exchanges and especially scientific work. The prototype at the European Organisation for Nuclear Research (CERN) which turned into the world web was already aimed at communication between scientists. More generally, the ICT revolution and the explosion of electronic networks have been acclaimed as abolishing distances and announcing 'the death of geography'.

2.2 Adverse Mechanisms

Actors are connected by proximity networks in various dimensions: geographical and geopolitical, cultural and linguistic, institutional, thematic. Although these networks stretch across borders, the nation is the locus where several types of proximity tend to be high simultaneously.

The first example is the strength of cultural and institutional linkages within nations. The nationalist resistance to globalisation of scientific communities, which peaked in the periods of world wars and also of cold war, is out of fashion, although the enrolment of science in strategic technology, not only military, is stronger than ever in this early XXIst

[1] On the scientific infra-structure see Irvine et al.(1997)

century. The national level has nevertheless been the main level of decision in the past and there are some clues that the inertia rooted in cultural traditions (a strong instance of proximity), specific mechanisms and political institutions make the 'National systems of innovation' (NSI, Lundvall, 1992, Nelson, 1993) more resistant than expected to the new momentum[2]. Even for the non-physical flows of knowledge, scientific communication, national borders still exist in many respects. The national level remains a major level of governance and funding; the institutional framework is still mainly national; the cultural and linguistic habits are also largely based on national specificity; multinational firms, as Pavitt and Patel have shown in several works (especially 1991), bring fewer internationalisation than expected: know-how is the less internationalised aspect of firms and most Multi-National Enterprises research remains firmly anchored in their home base — with perhaps signs of change in the recent period. National structures of Industrial Property Rights are also a resistant core (Foray, 1995). Internationalisation of systems of innovation has been discussed, for example, by Nosi and Bellon (1994), Carlsson (e.g. 1997, 2003), Archibugi et al. (1999). A grouping into families of NSI rooted in political and institutional heritage is found in Amable et al. (1997).

A second instance of a proximity based mechanism is the concentration and agglomeration processes at short distance. Complex short range relations between science, technology, industry, manpower and services, nourish spillovers and sustain local clusters, a new version of Marshallian districts (Beccatini, 1990) adapted to a knowledge based society, widely discussed in the economic literature. Particularly, the localised externalities from academic research have received much attention from scholars in the last decade (Jaffe, 1989; Audretsch and Feldman, 1996; Anselin et al., 1997). Proximity–sensitive exchanges of tacit knowledge are given a key role in these processes. Though based on codified publication, scientific communication does not escape the process, since science in action also requires exchanges of tacit knowledge exchange and fruitful face-to-face interactions (DeBeaver, op.cit.; Storper and Venables, 2004). At a wider scale, large regions have reinforced their co-ordination potential and funding capabilities. This may result in a changing prospect of world competition, where the 'regional system of innovation' (Cooke, 1998; Storper, 1997) as well as NSI compete in the knowledge–based markets.

Though particularly dense within a nation, cultural and linguistic proximity, as well as self-maintaining networks of sociability, shape preferential channels of communication at the international level.

[2] As discussed in NSI literature, these systems may not strictly coincide with national borders.

Knowledge does not travel as fast as information. Whilst some scholars anticipate a drastic reduction of ICT costs — including those of tacit knowledge exchanges — able to reduce the role of proximity (Foray and Mairesse, 2002), others (Morgan, 2001) strongly react at the thesis of the ‘death of geography’. The ability of ICT of getting rid of proximity effects and/or strong inertia of socio-political structures should not be overstated. Reshaping of communication networks will probably be slower than expected. A new impulse, somewhat paradoxically, may come from the regional systems of innovation, and especially from attractive high-tech districts which initially stemmed from short-range mechanisms, can reveal attractive for foreign actors and rich in long-range interactions. In the long term competition and cooperation amongst districts will perhaps erode the national borders and eventually turn into an internationalisation engine.

2.3 Some Internationalisation Measures Amenable to Bibliometrics

Internationalisation and internationalism in science take a variety of forms (Crawford et al., 1992; Elzinga and Landstrom, 1996). They encompass all dimensions of research systems: economic resources (programmes and funding systems; shared infrastructures, bilateral and multilateral agreements); human resources (teaching system and labour market of skilled manpower: PhD, postdoc, scientists; migrations, diasporas and networks, brain drain and brain gain); rules and norms of the community; general policy and governance levels.

In each area various modalities of internationalisation can be observed. One concerns the reduction of particular market imperfections owing to the national factors. Examples are progresses in international skilled labour mobility and reduction of nationally oriented publishing behaviour. Another axis concerns coordination and cooperation mechanisms, with sometimes a focus on reduction of international unevenness (EU structural funds and framework programs). Whether the reduction of barriers to competition and collaboration leads to a more equal distribution of final outcomes — the convergence question — is a crucial issue of globalisation studies. The question arises in a critical manner for the brain–drain, where internationalisation of skilled labour market has resulted so far in a strongly asymmetrical flow between the US and the rest of the world, with high benefits for the centre (Stephan and Levin, 1999). More generally, bibliometric distribution studies provided overwhelming evidence that scientific competition does produce skew distributions. Internationalisation is far from being a consistent process where removing barriers would necessarily mean a reduction of discrepancies.

In the following we will address three forms of internationalisation involving publications, and thus amenable to bibliometric measures at the macro-scale:

a) Internationalisation as a reduction of national barriers to competition: is scientific communication internationalised? We will focus on the core of 'certified' communication, scientific journals, which are a central locus of communication and competition among scientists.
b) Internationalisation as a reduction of national barriers to cooperation: it is generally admitted that the fabric of scientific interdependence networks, at the international level, is ever tighter. Does it mean a more open space? Here we will have a look at co-publication networks.
c) Internationalisation as a reduction of the national factor in final outcome distribution: are empirical convergence phenomena observed for the output of all (or groups of) countries? Convergent evolution and catch up processes are expected from targeted policies within supra-national economic communities (EU). We will report a few partial observations on this phenomenon.

3. IS SCIENTIFIC COMMUNICATION INTERNATIONALISED? THE CASE OF SCIENTIFIC JOURNALS

A basic fact about science is 'publish or perish'. Sociologists of science have devoted much effort to studying the role of publication in central aspects of self-organisation of the scientific community: circulation and archiving of information, priority issues, evaluation, etc. As a result the main media of communication, the scientific journal, has attracted many works, especially the impact factor issue (recent review by Glänzel and Moed, 2002). Rigidity and national enclosure in the main media of communication would mean a serious obstacle to internationalisation of science.

3.1 Marginal, Eroding but Still Alive: the National Model of Communication

The ideal type of 'nationally centred' model of science can be defined by the exclusive relation of domestic authored publication with domestic publishers and domestic language, symbolised by the prevalence of the 'nationally oriented' journals. Hence strong barriers to communication, competition and cooperation on three areas: among scientists; among publishers; among languages. The 'international' or 'trans-national model' assumes the disconnection between the three aspects (Zitt et al., 1998b): scientists compete for access to most visible media; publishers, either scientific societies or commercial publishers, try to push their influence by attracting visible authors; even languages compete for the largest international audience.

The long-term evolution since WWII of scientific communication in various disciplines can be seen as the transition from the first model to the second one. This national model has long lasted in countries such as the USSR, but also to a certain extent in certain disciplines of 'second–best' countries with strong editorial traditions, as the influential Garfield's diagnosis (1976) of the French situation demonstrated in the mid seventies. This transition process is largely advanced at the turn of the XIXth century, but the question can be extended to large emerging countries such as China. The competition game then redistributes roles and positions, not necessarily in the form of a more even distribution. For example, competition between languages has resulted in the quasi-monopoly of English as the *lingua franca* of primary communication, other languages being mostly confined to transfer purposes in particular geographical areas. The publishers' market is concentrated within operators, commercial publishers and/or societies, in a few countries (first of all the UK, the US, the Netherlands). Most publishers promote international journals, sometimes by merge between complementary national media, for example to form 'European journals'. Researchers tend to select a journal for their publications in terms of international visibility and citation rewards rather than national audience, as far as primary communication is concerned (transfer literature is another question, see also Chapter 20 by Lewison in this Handbook).

3.2 National Orientation of Journals: Static Measures

The international model predicts that journals, as spaces of competition for the authors, and themselves in competition, should increasingly welcome authors from various origins, and finally reflect the international profile of their scientific speciality in the world rather than their mother country's production. This deviation to the international profile of the discipline/speciality, used as the reference, operationalises 'relative internationalisation measures' of individual journals (Zitt and Bassecoulard, 1998a)[3]. A journal will be termed 'international' (static definition) if it reflects the national balance of the reference set at a given time. Many variants of internationalisation indices can be proposed: for example, by using a regional (geopolitical zones) breakdown instead of a national breakdown[4], by introducing a stratification by impact levels, by picking different statistical indices, relative or absolute[5]. Correlation of internationalisation indices with the journal impact is quite moderate (Bassecoulard and Zitt, 2004)[6]. These families of indicators can be extended to the study of the national profile of authors citing the journal, of authors cited by the journal, of editorial committees (studied for example by Braun and Bujdoso, 1983, see also Chapter 4 by Braun in this Handbook). Other measures, bibliometrics–based or not, include the scope of subscribers or readers (Wormell, 1998, Rey-Rocha and Martin-Sempere, 2004).

The distribution of relative internationalisation indices distribution for journals belonging to the *Science Citation Index* (SCI) or SCI–Expanded (Figure 18.1) suggests a mix of two populations, a majority class of international journals, and a small minority class of nationally oriented journals. The coexistence of two populations has some consequences in bibliometric comparative studies, briefly recalled below.

Longitudinal series of relative measures of national orientation based on deviations from an average world value are directly interpretable in terms of 'internationalisation' in a dynamic sense, i.e. a convergence process at the world level between journals authoring profiles. A clear upwards trend of internationalisation measures is observed in all disciplines of SCI (Zitt and

[3] Relative measurements are sensitive to artefacts in the delimitation of the reference, especially: the definition of the perimeter of specialities or disciplines; the coverage of the database. For example, strong biases in the coverage of SSCI and A&HCI in some disciplines prevent relative measurements, for lack of sound reference in these disciplines.

[4] For example, many European journals result from the merging of national journals from two or more countries (e.g. Astronomy and Astrophysics, European Physical Journal). Undoubtedly these were 'international journals' at the time they were created. Today, depending on the way the EU is considered (a mere club of countries or a real entity), a journal should probably be 'triadic' or multi-continental to be considered as truly international.

[5] Examples of the latter: number of distinct countries (of institutions) publishing in the journal; share of the country ranked number one -- or quantile-based share; concentration indices of authoring countries in the journal (for a review of the use of concentration indices in a bibliometric context see Egghe and Rousseau, 1990). The absence of a term of comparison in such indices can lead to undesirable results for journal assessment.

[6] Although there is a large overlap between top (respectively bottom) classes of impact and internationalisation. Moreover, the correlation between the level of internationalisation and the level of co-authorship (rate of co-authorship in the journal) is lower than +0.5 in most disciplines.

Bassecoulard, 1999a; Bassecoulard and Zitt, op.cit.), a trend not only owed to the erosion of the nationally oriented class which decreases but does not disappear.

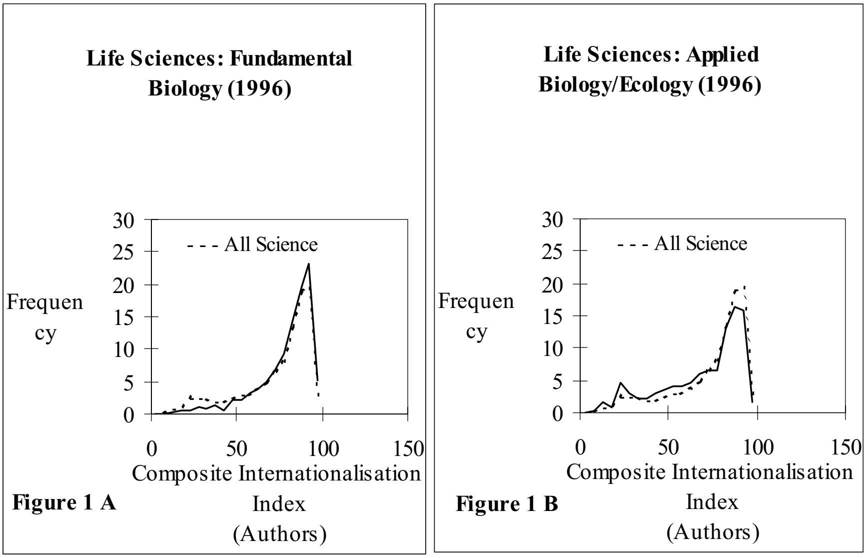


Figure 18.1. Distribution of journals by level of internationalisation (two disciplines).

The distribution shows a long tail or bi-modal shape, suggesting a mix of two populations, nationally oriented journals (minority) and a core of international journals. This finding is robust for a large variety of indices. Source: Z&B, ISI data (SCI), processing OST and INRA, first published in Scientometrics, 1999

An example of evolution is given Figure 18.2, for fundamental biology and applied biology. ISI keeps the perimeters of SCI or SCI–Expanded beyond the borderline of international journals. The survival of the nationally oriented category can be attributed to several factors: resistance of the 'national model' especially in (non English speaking) 'second best' countries with national editorial traditions; ISI policy towards emerging countries' promising journals while they still show little internationalisation; marginal generosity of ISI towards secondary communication. The figures of average deviations (variance as well as maximum deviation measures) confirm a steady trend towards internationalisation (*ibid.*).

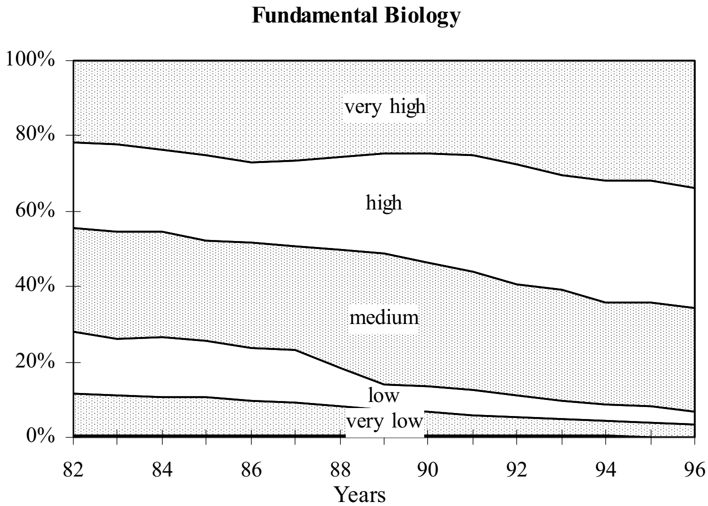


Figure 18.2a. Distribution of publications amongst journals by level of journal internationalisation: Fundamental Biology

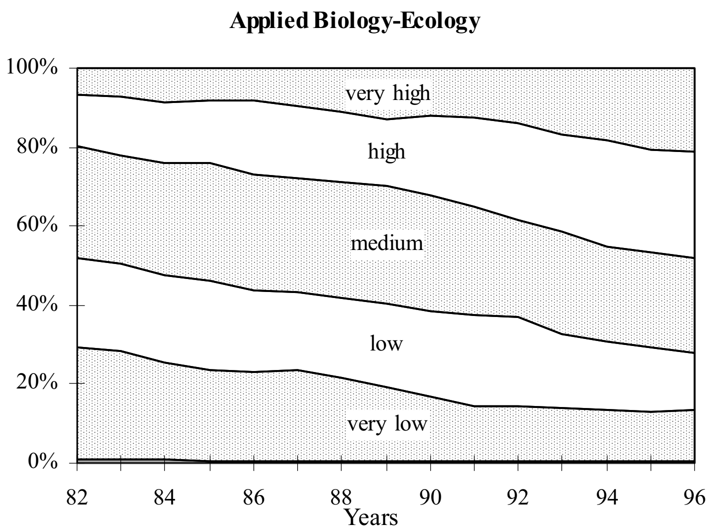


Figure 18.2b. Distribution of publications amongst journals by level of journal internationalisation: Applied Biology – Ecology. These figures illustrate the trend towards internationalisation, with strikingly contrasted profiles and top-classes gaining importance. Source: *ibid.*

3.3 Consequences for Interpretation of Bibliometric Indicators, Static and Dynamic: the Rent of Transition

Indicators are based on databases, and each database has its own statistical characteristics. It is generally considered that SCI (or now the '*Web of Science*') gives a good image of international science, based on a careful selection of journals, but is not bias–free. The interpretation of national bibliometric indicators, amongst other problems, should care for two related issues.

The first issue holds that the mix of two journal populations, with, for national oriented journals population, an uneven distribution among countries, has heavy consequences in the international benchmarking of outputs, including in static assessment. There is a serious risk of the overestimation of publication share and the underestimation of impact for the couples country–discipline, still marked by a national model reflected in SCI tail (Zitt et al., 2003). A similar effect for the presence of journals not using the English language was analysed in Van Leeuwen et al. (2001).

Discarding the national oriented class and the source of noise, allows one to correct series of classic indicators as a function of the database perimeter (*ibid.*) and to uncover nice bibliometric regularities which give another approach of international benchmarking. However, attention should be paid to applied field researches with possible national or regional specific targets (medicine, agriculture, etc.), with models that have been hardly received by the international literature, especially in developing areas. Secondly, in the case of countries in transition, emerging from a quasi-autarchic model, there may be a risk of discarding important media of primary communication.

The second issue holds that a related phenomenon, mentioned in Zitt et al. (1998b), affects the interpretation of longitudinal series of indicators. We can rephrase it as the 'rent of transition' for countries (or for disciplinary sub-system in countries) with strong scientific traditions, converting their scientific potential from national to international literature (adoption of English language, targeting of visible journals). All things equal, the consequence is a lasting upwards trend in the 'market share' of publications measured in the SCI. When the conversion to the international model is completed, the rent of transition disappears. For impact measures the case is much more complex.

It cannot be excluded that temporary or lasting decreasing return in 'visibility', measured by impact, is the price to be paid for a strong effort to increase volumes of publications in journals covered by ISI. We find again the trade-off between market share of publications and visibility (relative impact)[7]. These diminishing citation returns, in the long term, will halt if the benefits of worldwide competition extend to the newcomers. *Mutatis mutandis*, the case of swiftly catching up countries such as China is rather interesting to witness in this respect (see Chapter 22 by Jin and Rousseau in this Handbook).

When interpreting comparative long-term series, these mechanisms should be kept in mind. For example, an increasing trend in the world share of a particular country can be owed to actual progress of the research and innovation system (funding, capacity building, efficiency, etc.) in this country; to changes in communication strategy with a deliberate target on international media and language; to a particular policy of the database producers towards the country.

3.4 Communication in the Electronic Era

This chapter is focused on 'certified' media, and does not address the on-going revolution of scientific communication which is the object of a large literature. Let us only recall that in the electronic era, pre-prints, self-archives and other modes of quick communication, which pervade biology after physics, may alter the nature of the scientific article and the peer review process. Although a variety of alternative communication modes exist (books in some social sciences and Arts and Humanities, conference proceedings, with a low ratio of transformation into articles, in those disciplines and computer sciences), peer reviewed journals draw their legitimacy, as an intermediation, from the organisation of certification and archiving. For the first time, perhaps, anticipation of the mid-term future of the system becomes difficult, since alternative models can emerge both for certification and archives, on principles of self-organised and decentralised science (Ginsparg, 2000; Harnad, 2001). An indisputable progress in internationalisation of communication, besides the web posting of many types of scientific documents and teaching material, is the online availability of journal articles through electronic portals, which can be a bonanza for countries or provinces deprived of easy access to literature (see in particular

[7] Such trade offs are also observed when the behaviours of scientists change; for example under an external pressure. A quasi-experimental case has been recently discussed by Butler (2003, and her Chapter 17 in this Handbook), about the Australian policy of funding research in proportion of publications, with, as a result, a significant drop in impact.

the RFBR[8] initiative in Russia). The role of the Internet in access of peripheral countries to information and knowledge is a stake at the planetary level[9]. Free or easy access to many sources contributes to open competition, with obvious limits for tacit knowledge exchanges.

4. INTERDEPENDENCE NETWORKS

We will mainly focus on co-authorship networks, which represent an instance of collaboration strong enough to receive the sanction of the 'certified' literature — as mentioned above many other types of collaboration networks exist.

4.1 Co-authorship Networks

Owing to their richness of interpretation and their documentation at the institutional level in several databases, co-authorship networks have given birth to a huge number of contributions from the theoretical, methodological and political point of view. The reader is referred to Chapter 11 in this Handbook by Glänzel and Schubert for methodological points and a bibliography. In this section we will focus on some determinants and limits of international collaboration. Basically collaboration is driven by the same engines as other internationalisation mechanisms, in the framework of strong cultural and national constraints. The need for collaboration, the first engine of internationalisation, inherent in the scientific community, is anchored in the complementarity of competences. Collaboration is generally seen as a natural response to specialisation and increasing competition pressure, and brings better citation returns (Herbertz, 1995), even contributing to an inflation of citation figures (Persson et al., 2003). The term 'coopetition' was coined to reflect the mixes or changeovers of collaboration and competition. It applies quite well to scientists' behaviour. Top down initiatives back this trend for more collaboration, but sometimes take a form of coordination of large programs, leading to juxtaposed rather than co-authored articles/reports. Some tension may exist between bottom-up and top down processes (Ziman, 1994; Georghiu, 2001, with some special attention paid to European programs).

[8] Russian Foundation for Basic Research.

[9] The first phase of UN WSIS (the World Summit on the Information Society) was held in Geneva 10-12 December 2003, with a moderate success however.

4.2 The Evidence of Increasing International Collaboration

The evidence of a steady increase in international co-authorship has been stated by many scholars, on the basis of ISI data, using various counting methodologies (see the abovementioned chapter). The trend is quite strong, by and large the proportion of internationally co-authored papers is roughly doubled in a decade's span, 1990–2000 (OST figures on ISI data, ca. 7% in annual growth rate) without any apparent sign of saturation. Several remarks lead us to weigh this statement: first, not only foreign collaboration but all collaboration has developed in science, with a steady trend on bilateral and multilateral co-authorship; secondly, the 'within country' collaboration remains overwhelming in most large countries; thirdly, we observe a remarkable inertia of channels, which needs a few comments.

4.3 The Global Inertia of Channels

A fairly high contrast exists between the fast growing intensity of international collaboration flows and the relative inertia of collaborative channels. Complementing gross flows and the Salton measures (see Glänzel and Schubert's maps in Chapter 11 in this Handbook), size–normalised measures, especially the probabilistic affinity index or 'mutual preference' with appropriate setting[10], are particularly aimed at the detection of privileged channels, often mirroring cultural and geopolitical relations in a spectacular way. Although this type of index is extremely sensitive, it shows a remarkable stability, at least for large countries' pattern: flows keep swelling but in stable river beds. For example, over a decade, whilst the total intra-European co-authorship activity followed the world trend, the international preferences of France, Germany, and the UK remained relatively stable, with a strong cultural and historical (sometimes colonial) imprint, and this was also true for the USA and Japan (Zitt et al. 2000). These structures of co-authorship bring some evidence that cultural and geopolitical proximity — along with domination effects — supersedes

[10] Relative indices were advocated for example by Luukkonen et al. (1993). The index PAI= (n(..) × n(i,j)) / (n(i.) × n(.j)), on the contingency table of transaction, is the ratio of observed to expected flow. It needs some correction if one wants to ignore self co-authorship that inflate diagonals at the expense of other cells, yielding undesirable effects from skew distribution of actors: an iterative process of diagonal calculation towards the neutral value is recommended. A renormalisation of the interval is also useful. PAI-based networks, as well as gross flows, can be used as bases for various social network characterisation.

geographical proximity amongst infra-structural factors. Importance of the cultural factor was put forth in the earliest works on collaboration (De Beaver and Rosen, op.cit) and stressed by many authors (for example Okubo, 1996). Moreover, the voluntarist process at work in the European Union still seems far from bringing about an ineluctably homogeneous collaboration space, as shown with other methods by Leydesdorff, 2000, Grande and Peschke 1999 (for an earlier picture see Moed et al., 1991). However, some changes in affinity profiles of European peripheral countries is noted (Bassecoulard et al., 2001). Of course, in the long run geopolitical, if not cultural, relations rearrange networks. Political decisions or geopolitical earthquakes have transformed the historical affinity between the US and Japan, between the Western and the Eastern-Europe countries (Braun and Glänzel, 1996), or to a lesser extent between France and Russia or South America. But the stability in the medium run is quite remarkable. Infra-structural factors are a first natural explanation of the inertia of channels. The literature addressing the various determinants of collaboration flows usually retains geographical proximity, cultural/geopolitical proximity, inclusion in the same innovation system or nation. A detour by 'within countries' observation may be helpful.

4.4 A Regional Detour

Studies at the regional level within a country gave evidence of geographical proximity effects (Katz, 1993). Addressing international exchanges with a finer (infra-national) breakdown allows one to surmount a limitation of purely international measures in assessing the specific role of national borders. Their effect can be tested against a reference, namely, within-country regional borders. Studying the case of France and its neighbour countries, Okubo and Zitt (2004) show the overwhelming role of national borders, even in the case of border regions such as Alsace with bicultural traditions. The relative inertia of channels observed at the international level is also witnessed to a large extent at the inter-regional level. Within more closely connected countries (Scandinavia) cross-border regions with strong incentives, such as Oresund, may result in trans-border systems, but it is perhaps too early to assess such developments.

The regional detour corroborates the hypothesis that proximity factors underpinning collaboration networks rank as follows: institutional/national system (which also embody historical and cultural imprints); geopolitics and culture; geographic proximity. Other factors, less stable such as thematic alignment, also matter. Various factors have been combined in regression/gravity models (Nagpaul, 1999, see also in this Handbook Chapter 29 by Guellec and Van Pottelsberghe on technology).

A second explanation of stability should probably be sought at the individual level. The ‘quasi-neuronal’ persistence of inter-individual or inter-institutional linkages is a form of uncertainty reduction behaviour, with lasting linkages based on trust and maintained through face-to-face interactions in meetings and conferences. Combined with the infra-structural background, this factor could account for much of the relative inertia observed at the aggregate level, including at the international level.

4.5 Other Scientific Networks

Co-authorship is only a window on co-operation modes. For memory’s sake let us recall a few others also measurable by bibliometrics, either internal to science or hybrid, where the hypotheses of internationalisation can be tested: the network of scientific dependencies, as measured by citation flows between countries; the network of science–technology mutual interdependences (see the chapters and bibliographies on the Science–Technology Interface in this Handbook); the networks of Internet links, among them hyperlinks with the analogy ‘citation–sitation’ (see Chapter 15 by Ingwersen and Björneborn).

The study of linkages at individual and institutional level, especially, benefit from the ‘Social network’ toolbox borrowing from graph theory (amongst the early promoters Barnes, 1969), with recent developments (Watts and Strogatz, 1998; Zimmerman and Kirman, 2001). The social networks way of thought was also present in the pioneering works of ‘sociology of translation’ (early sketch of the actor–network theory, Callon et al., 1986, Turner et al., 1988). An example of application of social networks to technology is found in Chapter 28 by Breschi and Lissoni. These techniques are, for example, applied to co-authorship linkages (Erdös project among mathematicians; a growing number of works in physics literature), and bibliometricians are increasingly paying attention to social networks properties (Egghe and Rousseau, 2003).

5. GEOGRAPHICAL DISTRIBUTION OF KNOWLEDGE PRODUCTION: CONVERGENCE ISSUES

We have illustrated by a few examples the fading of national barriers in scientific competition and in gross cooperation flows (but with rather stable preferential channels). On the output side the acid test of internationalisation would be a more even distribution of knowledge production worldwide. This outcome is not precluded in internationalisation of 'coopetition' which could even lead to a reinforcement of inequalities and dominant positions. Do we witness a convergence in per capita scientific production? Do we witness a convergence in scientific specialisation? We limit ourselves to a few empirical indications on the movements within a decade.

5.1 International Concentration of Scientific Output

a) Big versus small scientific countries.

The simplest concentration indicators are the output shares of the first countries' decile(s). The top-decile accounts for 89% of output in 1991, dropping regularly to 85% in 2000/2001[11]. The second decile increases its share, from 8 to 11%, so does the third decile. A synthetic indicator of the cumulated distribution, the Gini index, also shows a slow and regular downwards trend of concentration (0.92 to 0.90), a trend confirmed by the coefficient of variation (standard deviation/mean).

Is this trend confirmed by citation distribution? We might observe an internationalisation of competition which eventually reinforces 'Matthew effect' and acquired positions; for example, conceding significant new publication markets for newcomers, but much smaller opportunities in the citation market still dominated by a few mainstream actors. Internationalisation has a completely different meaning if it covers an increasing concentration of scientific power or at the opposite if it yields, through transfers of competences, a more equal distribution of visibility. The figures confirm that concentration remains very high (still higher, as expected, than for publications), the ten major cited countries represented 95% in 1991 and 92% in 2000/2001, against a rather steady trend. Gini indices and CV confirm the very slow but real increase of evenness.

b) Mainstream versus emerging countries.

We may have a look at several sets of countries including more active or more productive countries. We paid attention to following perimeters: OECD, OECD plus countries with largest output (29 countries) noted OECD+, plus a tentative 'mainstream' perimeter (noted OECD–[12]).

[11] Source of indicators: INRA-Lereco; of output figures: OST, based on ISI ICF data.

[12] 'Mainstream' class has been defined as OECD, plus Israel, minus overlaps with an 'Emerging' class (Europe: accession and candidate countries, Turkey; Latin America: Mexico, Chile, Argentina, Brazil; Africa: South Africa; Asia: China, India, Taiwan, Singapore, South Korea). 'Peripheral' class groups other countries.

OECD (in its current perimeter, retropolated) represented 85% of world publication in 1991 vs. 83% in 2000/2001, slight contraction confirmed by citation shares (95%–93%).

'Mainstream' countries accounted for 83% of publications in 1991 and 79% in 2000/2001 (95%–92% of citations). The 'Emerging' class gathers most of the difference, jumping from 15% to 18% (citations 5%–8%). 'Periphery' remains marginal and stable.

Among a group of major countries (OECD current perimeter + countries with strongest output in the decade), concentration is falling more rapidly (Gini 0.71–0.65, citations 0.79–0.75). The picture is similar for OECD alone (0.73–0.68).

c) The EU.

If we turn now to the EU15, the world share is slightly growing (31 to 33%). Within EU15, Gini on publications looses four points over the period (0.58–0.54) and three points on citations (0.61–0.58).

d) Global picture.

There is a slight and regular drive toward reduction of concentration, both on publications (Figure 18.3) and citations. The slow contraction of mainstream, especially North America, is mostly captured by a class of emerging countries, especially in Asia, with the spectacular case of China, for example, rather than by periphery.

5.2 Convergence in Per Capita Publications (OECD Countries)

Per capita output data yield a complementary view. Using demographic series available for OECD we observed changes in per capita output, which can be held as a convenient basis for assessing convergences. Publication output growth rate (1998–2000 versus 1991–1993) decreases with level of output, as shown in Figure 18.4 that suggests catch up processes for smaller science producers. This is also true for the EU subgroup. Korea, Turkey, Portugal, Mexico, Greece and Poland have enjoyed important relative growth. In medium sized countries, Spain and Italy are on a remarkable upwards trend.

But the pace of catch up is slow. If the coefficient of variation is decreasing, the standard deviation amongst countries remains almost stable. Given the skew distributions, weighting by country size (output) does not allow a clear move to be recorded. The hierarchy of per capita scientific output in OECD, with Nordic countries, Switzerland, and the US ahead, is not likely to be deeply altered in a midterm future. Interchanges in ranks of per capita output are rare (Kendall tau > 0.87 for OECD, > 0.88 for EU15).

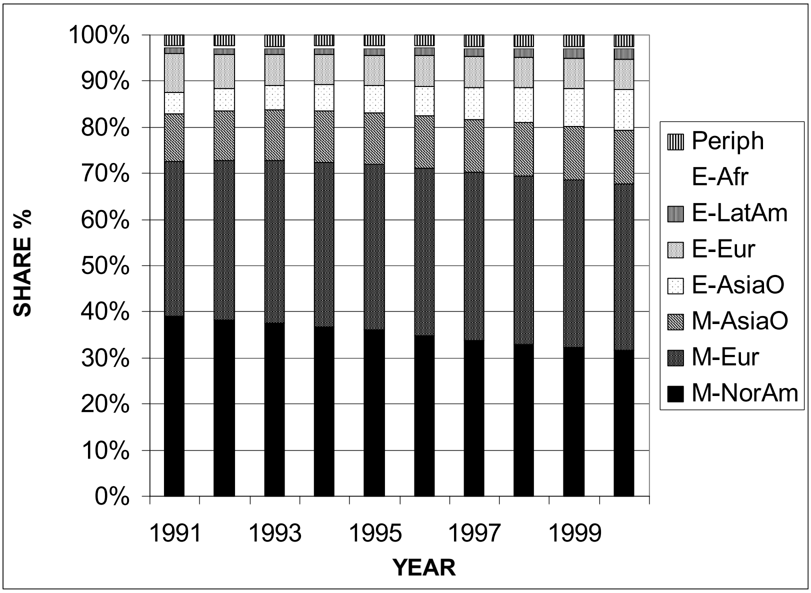


Figure 18.3. Distribution of publications by geopolitical area.
Source: ISI data, processing OST and INRA

The prefix M stands for mainstream, E for emerging. M-AsiaO comprises Japan, Australia, NewZealand. E-LatAm : Argentina, Brasil, Chile and Mexico. M-Eur: EU and northwestern Europe. E-Afr: South Africa. E-AsiaO is the great winner of the rearrangements. M-Eur is also expanding, US and Canada are on a downwards trend. Figures may be sensitive to artefacts in the coverage of the ISI database**.**

Evolution is similar in the EU15 sub-group: same relation between growth rates and output, evidence of decreasing non-weighted CV but imperceptible downwards trend of weighted indicators. These results tend to confirm our earlier observations (1999b). These data on per capita output tend to support the hypothesis of a slow move towards evenness. Again results of citation data confirm the slight progress in the reduction of inequalities.

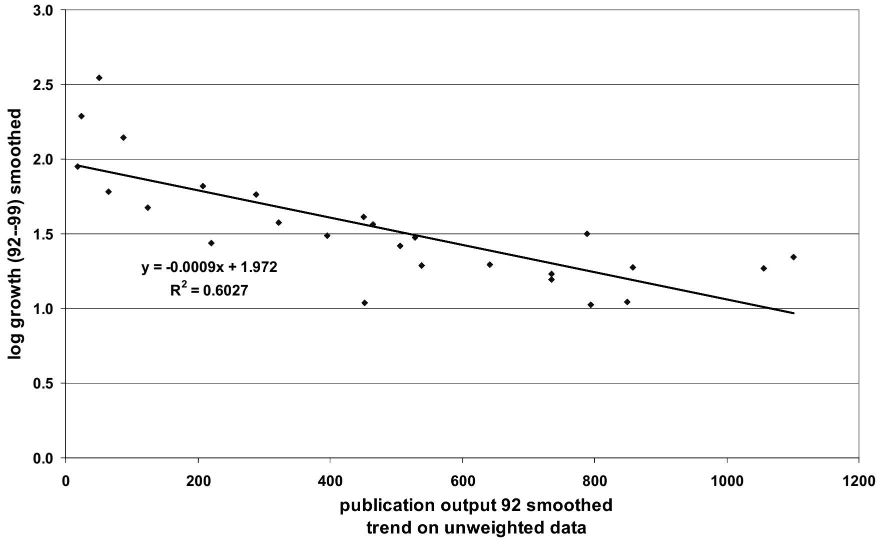


Figure 18.4. Growth versus Output.
Source: ISI data, processing OST and INRA

Not represented: USA, Canada, Czech Republic and Slovakia. World growth rate of publications cannot be interpreted as such, since it reflects the ISI database coverage policy. Only the country comparative trend should be considered, with respects to biases studied in literature.

5.3 Convergence in Thematic Specialisation

The third aspect of the homogeneity trend we consider here is the convergence of scientific specialisation. If internationalisation of output is assimilated to homogenisation at the world level, its empirical measure is by and large a reduction of discrepancies among countries, including in their specialisation patterns. Scientific specialisation is a complex phenomenon, linked to internal dynamic factors and public policy choices combined with agglomeration and learning processes. In some cases comparative advantages in terms of factor costs may also play a role, especially for developing countries, that can be restricted to disciplines requiring less funding and equipment. Analogies with international trade and patent economics in the explanation of specialisation should be carefully handled. The globalisation engine also conveys priorities external to academic science, towards profitable areas of technology and social needs, following a trend à la Schmookler, revised in more interacting fashion (Gibbons et al., 1994; Etzkowitz and Leydesdorff, 1997).

The configuration of science also evolves in deep movements, and the increasing role of ICT, and mainly biology, at the expense of physical sciences has been seen as featuring a 'new regime' for science, pioneered by most advanced countries especially the US. Bonaccorsi (2002), renewing Price's perspective, proposes a few characteristics of regimes and sees the compliance with the new trends as a key predictor of institutional success. A sketch of the new regime is found in Laredo (2002). Holding specialisation advantages in historic areas of specialisation or turning to new avenues is a crucial issue for policy makers. Despite a widely echoed internationalisation of priorities conveyed both by US policy and EU initiatives (biotechnology, nanotechnology, ICT), the worldwide convergence of specialisation is likely to be curbed by barriers to entry, irreversibility effects, and dynamics of geographical clustering.

Thematic convergence has been addressed in the literature, for example Doré et al. (1996). To sketch general orientations at the country level, it is convenient to aggregate academic disciplines into three large groups, respectively life/ physical/engineering. It should be noted that at the world level the balance between the groups, in the ISI database and the particular breakdown used[13], is fairly stable in the decade, ca 55% for life, 32% for physical, and 13% for engineering. Slight changes over the decade benefit the latter, a perhaps unexpected trend. It must be said that a representative balance is very difficult to achieve for database producers, including on theoretical grounds, and artefacts are unavoidable.

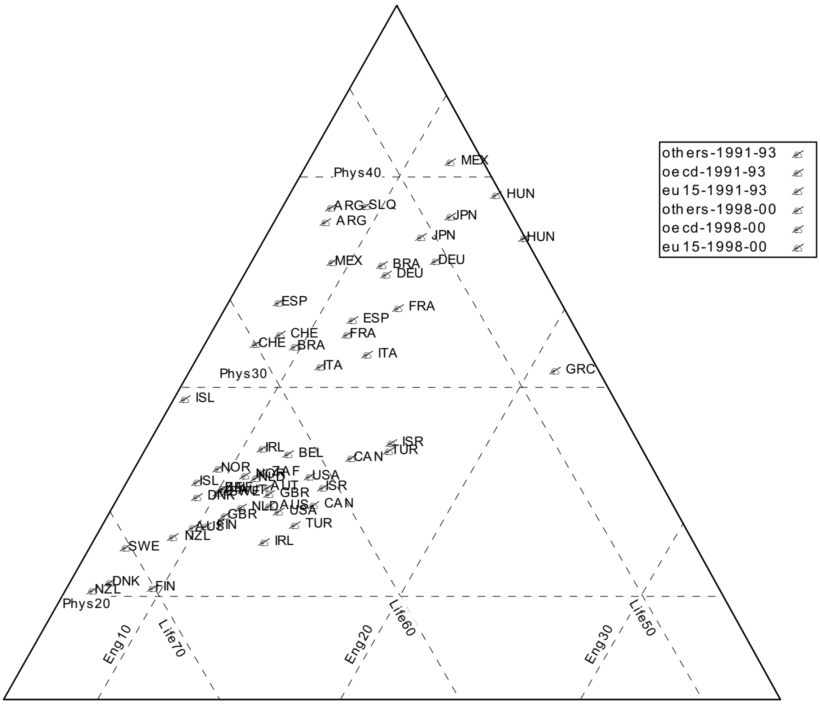
The balance for OECD and a set of selected countries is plotted in a triangular diagram Figure 18.5. Specialisation is very clear amongst actors, and clusters are relatively stable, few spectacular changes are recorded in the decade for large countries.

This relative stability is confirmed by quantitative measures. Two indicators, both based on discrepancies between country profiles and the world profile, were calculated: standard deviation on normalised activity indices (revealed advantage measure); quadratic distance to the average profile. Both were considered with and without country weights. Two disciplinary breakdown were used, the abovementioned three groups (D3), and the 8 academic disciplines (D8). The eight measures were calculated for each year.

At the world level no trend for homogeneity appears on the eight criteria, except moderately for variance in D8. All weighted measures (activity

[13] Life sciences entails medical research; fundamental biology; applied biology. Physical sciences: physics; chemistry; earth and space. Engineering sciences: engineering; computer science; mathematics has been joined to this group.

indices and inertia) revealed a slight divergence. On the perimeters including major countries (OECD, OECD+), convergence is noted on non-weighted indices and divergence on weighted indices, which suggests that important actors do not get closer.



Source: Z&B, 1999 - ISI data, processing OST and INRA. Graph by Tri-Draw software.

Figure 18.5. Disciplinary balance (1998 – 2000) vs. (1991 – 1993) – percentages ISI data, processing OST and INRA

This triangular diagram shows the shares of the three disciplinary groups (by country) summing at 100% as a result of fractional counting. Each country is represented at two periods. OECD countries and a few selected others are shown. For legibility only the bottom of the diagram is represented. Within the diagram two clusters appear: mainstream in the lower triangle, with Nordic and Anglo-American countries, biology oriented; large European and Latin America countries in the upper triangle: Life 55%, Physics 35%, Engineering 10%. Some movements are noted but the clusters remain stable. Outsiders (not shown): on the upper right of the diagram (much less than 50% in Life sciences), some countries of Eastern tradition keep the strong traditional involvement in physical sciences, between 50 and 70% and often less than 20% in Life Sciences (Poland, Russia, Ukraine, China); Korea and Taiwan, remote outsiders, have exceptionally high involvement in Engineering, ca. 20–30% and less than 35% in Life Sciences.

In contrast, the EU15 perimeter records a convergence on all measures, except weighted activity index which is fairly stable. Although these results should be confirmed on longer series and finer disciplinary breakdowns, a slow homogenisation process seems to be at work in Europe.

5.4 Convergence at the Regional Level

Let us focus on the European landscape. As seen above, inter-country unevenness is quite large for science, and this is also true for technology as measured by patents. But compared with technology, overall territorial inequality in science output has a strong regional component, owing to the scattering (and size – performance variability) of universities; in contrast, unevenness in patent output relies more on country international differences. There is some evidence (Zitt et al., op. cit. 1999b) that as far as science is concerned, regional inequality over EU15 is also on a downwards trend, but the landscape can differ among disciplines. Numerous works and reports address the stakes of EU convergence in STI issues (see for example Denozios, 1997) and the issue is topical within the new EU25. The above mentioned literature on spillovers has examined many sectors and case studies in regions. Understanding of dynamic phenomena of regional S&T clustering is a wide area for future research.

6. CONCLUSION

We have described various aspects of scientific internationalisation amenable to bibliometric measures based on published outputs. There is much evidence that internationalisation is growing but with contrasted facets.

Some barriers to international competition are being lifted. The national model in scientific communication is gradually being limited to secondary (transfer) communication, and scientific journals tend more and more to reflect the international variety of contributions in their discipline and level. Transition mechanisms between the national and the trans-national model, as well as the persistence of national media, should be taken into account for the interpretation of bibliometric time series. It should be stressed, however, that the disappearance of some market rigidity does not mean perfect competition, nor does it imply a trend toward more evenness. Dominant positions in editorial committees for example can still convey national power of mainstream countries.

Cooperation, coordination and interdependence are another target of internationalisation processes in science. In co-authorship relations, we observed an apparent paradox. On the one hand, gross flows show the most impressive changes amongst all other manifestations of internationalisation, but on the other hand changes occurring in the topography of preferential collaboration channels are rather slow. This relative inertia can be attributed to stable infrastructural factors, as well as feedback loops on existing individual relations. The landscape of collaboration draws more a 'network' rather than a homogeneous 'space'. The national and cultural barriers are resistant. Even in an activity where exchanges are mostly non-physical, geography is far from dead. For example, the degree of EU integration on this criterion has not followed the political impulse (Head and Mayer, 2000, show the same findings on intra-EU commercial exchanges). From the methodological point of view, progresses are expected from the new tools of social networks theory which could help to bridge micro and macro-approaches of scientific networks.

Turning to the world distribution of scientific output, the geographical distribution of knowledge production shows a decline of concentration, but at a very slow pace, in the universes considered (World, OECD, EU15). The evolution of concentration of output and citations on the one hand, the convergence in per capita publication and citation on the other suggest that the picture of world science production is slowly becoming less unequal. Whilst 'emerging countries', especially in Asia and also in Europe, are on a catch up trajectory in the latter decades, the periphery does not participate in the movement. The pressure of newcomers mechanically shrinks the relative share of dominant countries in the scientific communication, but to a very moderate extent. If the case of China is spectacular, relative rankings of OECD countries in per capita output have changed little in the decade. The other major phenomenon, the intensive draining of human resources by the US, also limits the long term prospect of convergence. The thematic specialisation, measured at the level of discipline, does not give evidence of a convergence process. Scientific specialisation, rooted in historical trajectories of NSI, resists, except a slow homogenisation process within EU15. Let us conclude with a few interrogations.

Interaction of internationalisation modes

The above perspectives on internationalisation are not independent. Collaboration as a merging of complementary skills can be interpreted as a response to the diversity of subjects and specialisation — in addition to rewards in terms of visibility. The relation of international collaboration, output growth, and geographical distribution of output is complex. Large countries offer a variety of in house collaboration targets so that they can afford low levels of foreign linkages (USA, Japan). At the opposite end peripheral countries exhibit very high rates of international collaboration, as a response to scarcity of local resources (for the African case Gaillard et al., 2002). Emerging countries use abroad linkages in catch up processes, but at the same time collaboration within the mainstream helps to keep high standards and advance. International collaboration also has ambivalent relations with scientific manpower migration, of both substituability and complementarity. Circulation of students and scholars, probably more than collaboration, and the related brain drain/brain gain balance, determine the dynamics of catching up.

New barriers to communication due to appropriation of science?

A particular concern is the connexion between science and technology internationalisation. There is growing evidence that the frontier between science as a public good and technology as a private good is becoming fuzzy, especially in the area of biotechnology and new ICT. The academic model of free communication can be threatened in various ways by the pressure of property rights (Nelson, 2004). If globalisation fosters appropriated forms of knowledge, it can delay exchanges or restrict their content. The biotech area exemplifies this new pressure on the traditional model of science.

For example, we have watched the slowly and regularly decreasing world share of the US in articles' output (at a much lower pace for citations). At the same time the share of the US in patents, including the European or PCT, is steadily increasing, without mentioning the defense area. This leaves some interpretations open: is it a simple consequence of emerging countries' differential pressures in basic and applied research? Or the consequence of a competition publication – IPR in knowledge diffusion, watched at the university level (Dasgupta and David, op.cit., Mowery and Ziedonis, 2002) and decreasing incentives to publish open science in some areas?

Lon-term dynamics: geopolitics and scientific regimes

Internationalisation has to be placed in the evolution of innovation systems and, in the long run, in the perspective of geopolitics. We have emphasised the role of infra-structural factors in shaping scientific collaboration networks, factors responsible for a relative inertia in the medium run but submitted in the long period, through the geopolitical component, to drastic changes. The transition to open political and/or economic systems (case of Spain and Portugal in the seventies, more recently of former Eastern block countries, of dragons and China, etc.) has deeply contributed to the competitiveness and sometimes to the emergence of scientific communities in these countries. The supra-national policy of the EU, first through structural funds, then through Framework Programs, probably explains why EU countries tend, albeit very slowly, to converge.

The EU is also a laboratory where supra-national, national and regional levels compete and complement each other in the shaping of the research and innovation system. Perhaps the most appealing question is whether the new 'regime' in the leading edge of science – biotech, information, nano – with agglomeration and coopetition amongst science districts, beyond national borders, will be able to destabilise the factors of inertia rooted in history and culture.

Internationalisation of topics

The drifts of nationalism and ideology in science dramatically curbed scientific exchanges during the XXth century. The Republic of Science wishes to ignore borderlines, but at the same time elitism and concentration are consubstantial with the community's norms and habits, expressed in skew distributions of output and 'Matthew effect'. As we have stressed, internationalisation of competition or cooperation does not promise a fading of borderlines in productivity maps. Neither do they warrant that variety will be safeguarded, especially in terms of heterodox thought and research topics. A particular question is about topics specific to developing/emerging countries, which may not find an echo in the international community. 'Nationalism in science' which found some prestigious advocates, for example Raman[14] in India, in the past century, can be seen as a refuge for addressing domestic issues (Arunachalam, 1997). The thematic orientations of domestic research, international research on the country's specific topics, and diaspora have been found very different in the case of an ultra-peripheral country (Bassecoulard et al., 2003). The marginalisation of periphery's preoccupations in agriculture, biology, and medicine would be a failure of internationalisation, which on other aspects brings hope for scientists and students to be able to access information from everywhere.

Whilst restrictions on international communication, competition, collaboration – and skilled manpower circulation – tend to fade, infrastructural factors, proximity effects, inertia of networks constrain the rearrangements. Like other globalisation processes, internationalisation in science is a *Janus Bifrons*, conveying antagonistic forces: on the one hand, through the reinforcement of (imperfect) competition and the Matthew effect, it may secure or enhance dominant positions; on the other hand, actors in transition or in emergence benefit from the circulation of information and skilled manpower. The empirical evidence is in favour of more evenness, but the trend is quite moderate. In the next decade one will observe whether the new regime in the leading edge of science is able to

[14] Sir C.V. Raman, Nobel Prize for physics (1930).

shake factors of inertia and to challenge — or reinforce — international inequality in the production of knowledge.

REFERENCES

Amable, B., Barré, R., Boyer, R. (1997). *Les systèmes d'innovation à l'ere de la globalisation*. Paris: Economica.

Anselin, L., Varga, A., Acs, Z. (1997). Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics,* 42, 442–448.

Archibugi, D., Howells, J., Michie, J., (Eds.). (1999). *Innovation policy in a global economy*. Cambridge Un. Press.

Arunachalam S. (1997). How relevant is medical research done in India? A study based on Medline. *Current Science*, 72 (12), 912–922.

Audretsch, D.B., Feldman, M.P. (1996). R&D Spillovers and the geography of innovation and production. *The American Economic Review,* 86 (3), 630–640.

Barnes, J.A. (1969). *Networks and political process*. In J.C. Mitchell (Ed.), Social networks in urban situations (pp.51–76). Manchester Un. Press.

Bassecoulard, E., Okubo, Y., Zitt, M. (2001). *Insights in determinants of international scientific cooperation*. In F. Haveman., R. Wagner–Döbler, H. Kretschmer (Eds.) , Collaboration in Science and Technology. Berlin: Gesellschaft für Wissenschaftsforschung.

Bassecoulard, E., Ramanana-Rahary, S., Zitt, M. (2003). *Ultra-periphery of science: three contrasting views of the Malagasy contribution – in terms of domestic research, the diaspora, specific topics*. In Proceedings of the 9th International Conference on Scientometrics & Informetrics. Beijing.

Bassecoulard, E., Zitt, M. (2004). *Scientific communication: two aspects of globalisation at the journal level*. In H. Laitko, A. Scharnhorst (Eds.), Festschrift für M. Bonitz, *forthcoming*.

Beccatini, G. (1990). *The Marshallian industrial district as a socio-economic notion*. In F. Pyke, G. Beccatini, W. Sengenberger (Eds.), Industrial Districts and Interfirm Cooperation in Italy. GENEVA: Int. Inst. for Labour Studies.

Bonaccorsi, A. (2002). *Matching properties. Research regimes and institutional systems in science*. Conf. Science as an institution, the institution of science. Siena, Jan. 25–26.

Braun, T., Budjoso, E. (1983). Gatekeeping patterns in the publication of analytical chemistry research. *Talanta,* 30 (3*),* 161–167.

Braun, T, Glaenzel, W. (1996). International Collaboration: Will it be keeping alive East European Research? *Scientometrics,* 36 (2), 247–254.

Butler, L. (2003). Explaining Australia's increased share of ISI publications – the effects of a funding formula based on publication counts, *Research Policy* 32, 143–155.

Callon, M., Law, J., Rip, A. (1986). *Qualitative Scientometrics*. In M. Callon, J. Law, A. Rip, (Eds.), Mapping the Dynamics of Science and Technology (pp. 107–123). London: Macmillan.

Callon, M. (1994). Is science a public good? *Science Technology and Human Values*, 19, 395–424.

Carlsson, B. (1997). *Technological systems and industrial dynamics*. Boston (MA): Kluwer.

Carlsson, B. (2003). *Internationalisation of innovation systems: a survey of the literature*. Conf. in honour of K. Pavitt, What do we know about innovation? SPRU, 13–15 November 2003.

Cooke, P. (1998). *Introduction, Origins of the Concept*. In H.J. Braczyk, P. Cooke, M. Heidenreich (Eds.), Regional Innovation Systems: the Role of Governances in a Globalized World. London: UCL Press.

Crawford, E. (1992). *Nationalism and Internationalism in science, 1880–1939, Four studies of the Nobel population*. Cambridge University Press.

Crawford, E., Shinn, T., Shörlin, S. (Eds.). (1992). *Denationalizing science: the contexts of international scientific practice*. Kluwer.

Dasgupta, P., David, P.A. (1994). Toward a new economics of science, *Research Policy,* 23, 487–521.

David, P.A., Foray, D. (1995). Accessing and expanding the science and technology knowledge base. *STI Review,* 16, 13–68.

De Beaver D., Rosen, R. (1978). Studies in scientific collaboration Part I. *Scientometrics,* 1, 65–84.

Deniozos, D. (1997). Relevance of research and technological activities for economic development in some less-favoured European countries. *Science and Public Policy,* 24 (3), 183–188.

Doré, J.C., Ojasoo, T., Okubo, Y., Dudognon, G., Durand, T., Miquel, J.F. (1996). Correspondence Factorial Analysis of the Publication Patterns of 48 Countries over the Period 1981–1992. *Journal of the American Society for Information Science,* 47 (8), 588–602.

Egghe, L., Rousseau, R. (1990). *Introduction to Informetrics*. Amsterdam: Elsevier.

Egghe, L., Rousseau, R. (2003). BRS–compactness in networks: theoretical considerations related to cohesion in citations graphs, collaboration networks and the Internet. *Mathematical and Computer Modelling,* 37, 879–899.

Elzinga, A., Landstrom, K. (Eds.). (1996). Modes of internationalism and science. London: Taylor Graham.

Etzkowitz, H., Leydesdorff, L. (1997). *Universities and the global knowledge economy: a triple helix of university–industry–government relations*. London: Pinter.

Foray, D. (1995). *The economics of intellectual property rights and systems of innovation*. In J. Hagedoorn (Ed.), Technical change and the world economy: convergence and divergence in technology strategies. Aldershot (UK), Brookfield (USA): Edward Elgar.

Foray, D., Mairesse, J. (2002). *The Knowledge Dilemma and the geography of innovation*. In M.P. Feldman., N. Massard (Eds.), Institutions and systems in the geography of innovation (pp. 35–54). Kluwer.

Gaillard J., Hassan M., Waast R., et al. (2002). *Africa in World Science*. Paris: UNESCO

Garfield, E. (1976). La science française est-elle trop provinciale ? (Is French research too provincial?). *La Recherche* 70, 757–760.

Georghiou, L. (2001). Evolving frameworks for European collaboration in science and technology, *Research Policy* 30, 891–903.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M. (1994). *The new production of knowledge: the dynamics of science and research in contemporary societies*. Sage: London.

Ginsparg, P. (2000). *Creating a global knowledge network: don't just clone the paper methodology.* Freedom of information Conf. New York Acad. of Medicine, 6–7 July 2000 (www.biomedcentral.com/info/ginsparg-ed.asp).

Glaenzel, W., Moed, H.F. (2002). Journal impact measures in bibliometric research. *Scientometrics,* 53 (2), 171–193.

Grande, E., Peschke, A. (1999). Transnational cooperation and policy networks in European science policy–making. *Research Policy,* 28, 43–61.

Harnad S. (2001). The self-archiving initiative. *Nature* 410, 1024–1025.

Head K., Mayer T. (2000). Non-Europe: The magnitude and causes of market fragmentation in the EU. *Weltwirtschaftliches Archiv/Review of World Economics*, 136, 284–314.

Herbertz, H. (1995). Does it pay to cooperate? A bibliometric case study in molecular biology. *Scientometrics,* 33, 117–122.

Irvine, J., Martin, B.R., Griffiths, D., Gathier, R. (1997). *Equipping science for the 21th century*. Edward Elgar.

Jaffe, A. (1989). Real effects of academic research. *American Economic Review*, 79, 957–970.

Katz, S.J. (1993). Geographical proximity and scientific collaboration. *Scientometrics,* 31 (1), 31–34.

Laredo, P. (2002). *Six major challenges facing public intervention in higher education, science, technology and innovation*. Triple-Helix Conference, Copenhagen, November 6–9.

Leydesdorff, L. (2000). Is the European union becoming a single publication system? *Scientometrics,* 47 (2), 265–280.

Lundvall, B.A. (Ed.). (1992). *National systems of innovation. Towards a theory of innovation and interactive learning*. London: Pinter.

Luukkonen T., Tijssen R.J.W., Persson O., Siversten G. (1993). The Measurement of international scientific collaboration. *Scientometrics* 28 (1), 15–36.

Merton, R.K. (1973). *The sociology of science. Theoretical and empirical investigations.* (Collection of works). Chicago: Un. of Chicago Press.

Moed, H., De Bruin, R.E., Nederhof, A.J. Tijssen, R.J.W. (1991). International scientific cooperation and awareness within the European community: problems and perspectives. *Scientometrics,* 21 (3), 291–311.

Morgan, K. (2001). *The exaggerated death of geography: localised learning innovation and uneven development.* The future of innovation studies conf., Eindhoven, 20–23 Sept. 2001.

Mowery, D.C., Ziedonis, A.A. (2002). Academic patent quality and quantity before and after the Bayh–Dole act in the United States. *Research Policy* 31, 399–418.

Nagpaul, P.S. (1999). *Exploring a pseudo–regression model of transnational cooperation in science*. 7th ISSI Intl Conf., Colima, Mexico, 375–385.

Nelson, R.R. (1993). *National innovation systems. A comparative analysis*. New York: Oxford University Press.

Nelson, R.R. (2004). The market economy and the scientific commons. *Research Policy* 33, 455–471.

Niosi, J., Bellon, B. (1994). The global interdependence of national innovations systems. *Technology in Society,* 16, 173–197.

Okubo Y. (1996). L'internationalisation de la science (Internationalisation of science). *Futuribles,* 210, 43–56.

Okubo, Y., Zitt, M. (2004). Searching for research–Europe: a closer look at international and inter-regional collaboration in the case of France. *Science & Public Policy*, forthcoming.

Pavitt, K., Patel P. (1991). Large firms in the production of the world's technology: an important case of non-globalisation. *Journal of International Business Studies,* 22, 1–21.

Persson, O., Glaenzel, W., Danell, R. (2003). *Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies*.

Proceedings of the 9th International Conference on Scientometrics & Informetrics Beijing, August 2003.

Price, D.J. de Solla (1963). *Little science, big science*. New York: Columbia Un. Press.

Rey-Rocha, J., Martin-Sempere, M.J. (2004). Patterns of the foreign contributions in some domestic vs. international journals on Earth Sciences. *Scientometrics,* 59 (1), 95–115.

Stephan, P.E. (1996). The economics of science. *Journal of Economic Literature,* 34, 1199–1235.

Stephan, P.E., Levin, S.G. (1999). Are the foreign–born a source of strength for US science? *Science*, 1213–1214.

Storper, M. (1997). *The regional world: territorial development in a global economy*. New York: Guilford.

Storper, M, Venables, A.J. (2004). Buzz: Face-to-face contact and the economy. *Journal of Economic Geography*, forthcoming.

Turner, W.A., Chartron G., Laville F., Michelet B. (1988). *Packaging information for peer review: new co-word analysis techniques.* In A.F.J. van Raan (Ed.), Handbook of quantitative studies of science and technology (pp. 291–323). Amsterdam: Elsevier Science Publishers.

Van Leeuwen, T.N., Moed, H.F., Tijssen, R.J.W., Visser, M.S., Van Raan, A.F.J. (2001). Language biases in the coverage of Science Citation Index and its consequences for international comparison of national research performance. *Scientometrics,* 51 (1), 335–346.

Watts, D.J., Strogatz, S.H. (1998). Collective dynamics of 'small–world' networks. *Nature,* 393, 440–442.

Wormell, I. (1998). Informetric analysis of the international impact of scientific journals: how international are the international journals? *Journal of Documentation,* 54 (5), 584–605.

Ziman, J. (1994). *Prometheus bound – science in a dynamic steady state*. Cambridge: Cambridge Un. Press.

Zimmermann, J.B., Kirman, A. (Eds.). (2001). Economics with heterogeneous interacting agents. *Lecture notes in Economics and Mathematical Systems*, 503. Springer-Verlag.

Zitt, M., Bassecoulard, E. (1998a). Internationalization of scientific journals: a measurement based on publication and citation scope. *Scientometrics*, 41 (1–2), 255–271.

Zitt, M. Perrot, F., Barré, R. (1998b). Transition from national to transnational model and related measures of countries performance. *Journal of the American Society for Information Science,* 49 (1), 30–42.

Zitt, M., Bassecoulard, E. (1999a). Internationalization of communication: a view on the evolution of scientific journals. *Scientometrics*, 46 (3), 669–685.

Zitt, M., Barré, R., Sigogneau, A., Laville, F. (1999b). Territorial concentration and evolution of science and technology activities in the European Union: a descriptive analysis. *Research Policy,* 28 (5), 545–562.

Zitt, M., Okubo, Y., Bassecoulard, E. (2000). Shadows of the past in international cooperation: Collaboration profiles of the top five producers of science. *Scientometrics,* 47(3), 627–657.

Zitt, M., Ramanana, S., Bassecoulard, E. (2003). Correcting glasses help fair comparisons in international science landscape: country indicators as a function of ISI database delineation. *Scientometrics,* 56 (2), 259–282.

Chapter 19

ANALYSIS OF CROSS-DISCIPLINARY RESEARCH THROUGH BIBLIOMETRIC TOOLS

María Bordons, Fernanda Morillo, and Isabel Gómez

Centro de Información y Documentación Científica (CINDOC), Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain. E-mail: mbordons@cindoc.csic.es

Abstract: A review of interdisciplinarity in science is presented from the point of view of quantitative studies of science. The main objectives pursued and methodologies used in publications on cross-disciplinary research are pointed out, as well as the most relevant results obtained. The study of cross-disciplinary collaboration between authors, co-classification analysis, interdisciplinary nature of publication journals and cross-disciplinary references and/or citations are the most useful approaches to the topic. Results about a global analysis of scientific areas and disciplines based on ISI multi-assignation indicators are presented.

1. INTRODUCTION

Disciplines are the intellectual and social structures through which modern knowledge is organised. However, the established order of knowledge is being influenced nowadays by an important change: boundaries separating disciplines are dissolving, disciplines tend to overlap, and new hybrid fields emerge. As disciplines are becoming more diffuse the number of new fields and specialties is growing and interdisciplinarity becomes a common experience. This was explained by Gibbons et al. (1994) as the emergence of a new mode of knowledge production: the 'interdisciplinary science', which coexists with the traditional 'disciplinary science', the former being reinforced by application oriented research. In fact, problem driven research is an important source of interdisciplinarity, since many of society's major problems, such as environmental issues, require integrating approaches from different disciplines. The increasing

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 437-456.

specialization in science is also a factor contributing to interdisciplinarity, because the combination of knowledge from different fields is necessary to cope with specific scientific issues, either practical or theoretical, internally or externally generated.

Cross-fertilization across different disciplines is being described as a key element in the advancement of science. Interdisciplinarity is being associated with high levels of creativity, progress, and innovation because many of the intellectual ‘breakthroughs’ of modern times were obtained by crossing disciplinary boundaries. However, different organisational and cognitive problems make the development of cross-disciplinary research particularly difficult because it requires extensive networks, considerable time, and researchers′ mobility amongst disciplines. In the most advanced countries governments are involved in the project of promoting cross-disciplinary research. They try to encourage contacts between disciplines, detect which are the requirements for high quality cross-disciplinarity research and adapt assessment processes to the new structure of knowledge in these borderline fields.

Different types of cross-disciplinary research have been described, as well as a specific terminology, unfortunately not shared by all authors. The most commonly accepted definitions come from the OECD (1998), in which multi-disciplinarity, interdisciplinarity and trans-disciplinarity are used to refer to increasing levels of interaction and integration among disciplines. In ‘multi-disciplinary research’, the subject under study is approached from different angles using different disciplinary perspectives, but integration is not accomplished. ‘Interdisciplinary research’ leads to the creation of a theoretical, conceptual, and methodological identity, so more coherent and integrated results are obtained. Finally, ‘trans-disciplinarity’ goes one step further: it refers to a process in which convergence between disciplines is pursued, and it is accompanied by a mutual integration of disciplinary epistemologies (Van den Besselaar and Heimeriks, 2001). In our study, we use ‘interdisciplinarity’ (ID) and ‘cross-disciplinary’ research as covering all those types of research described above.

As a sign of the increasing role of interdisciplinarity in science, we can mention the outstanding growth of the terms ‘interdisciplinarity’ and ‘multi-disciplinarity’ in scientific literature. According to a study by Braun and Schubert (2003), the growth of these terms in the titles of papers covered by the database *Science Citation Index* during the years 1980–1999 was exponential, with a doubling time of 7 years, much quicker than for science journals (15–20 years).

2. OBJECTIVES OF THE STUDIES

Nowadays the study of interdisciplinarity is addressed from different disciplines, such as Sociology or Philosophy of Science, but in this paper it will be analysed from the point of view of Quantitative Studies of Science. Different aspects of interdisciplinarity can be studied through bibliometric analysis, but answering the following questions are the objectives most frequently pursued:

1. Is there an increasing trend towards interdisciplinarity in modern science? The quantification of interdisciplinarity in science as well as its evolution over the years emerges as an objective in different studies. Different methods and indicators addressing these issues have been introduced during the last years (see for example Van Leeuwen and Tijssen, 2000; Braun and Schubert, 2003).
2. Are there differences between research fields as to their interdisciplinary nature? And more specifically, are there differences in interdisciplinarity of fields according to their basic/applied or to their hard/soft nature? (see for example Hargens, 1986; Morillo et al., 2003). It is assumed that the different disciplines are differently involved in cross-disciplinary activities and the sensitivity of several bibliometric indicators to discriminate between disciplines according to their degree of interdisciplinarity is tested in the literature.
3. How can we study the structure of science? We would like to study the structure and dynamics of science with special emphasis in the relationship between disciplines. Flows of knowledge between disciplines are analysed at the macro level by means of citation flows (see for example, Small, 1999; Van Leeuwen and Tijssen, 2000; Rinia et al., 2002; NSB, 2000); through migration of scholars between fields (Le Pair, 1980; Hargens, 1986); identifying 'boundary crossing' authors and groups (Pierce, 1999), or through the analysis of the relationship between disciplines by means of multi-assigned journals (Katz and Hicks, 1995; Morillo et al., 2001, 2003). Intellectual imports and exports from one discipline to another are identified and disciplines are classified accordingly.
4. Has ID a positive effect on research? Is it possible to confirm the benefits of cross-disciplinary research by means of bibliometric based indicators? Is interdisciplinary research of higher quality than single based research? Assuming that interdisciplinary research is particularly relevant for the advancement of science, cross-disciplinary contacts are fostered by government agencies. But is it possible to monitor the efficacy of science policy measures oriented to foster interdisciplinary research through

bibliometrics (for example Bordons et al 1999)? The higher quality of interdisciplinary research as compared to the single disciplinary kind can be analysed through citation counts (for example Qiu, 1992; Steele and Stier, 2000) and through the success rate in research grant applications or acceptance rate in scholarly journals.

5. Should interdisciplinary research be reviewed in the same way as conventional disciplinary research? Claims have emerged from different sectors stating that traditional methods used for the evaluation of research might be inadequate for the assessment of ID research. There is some evidence that cross-disciplinary research may be undervalued in the review process, owing in part to cognitive reasons, such as the intrinsic difficulties in assessing the quality of this type of research, but also to organisational factors, such as the fact that evaluation committees are usually organised into discipline categories and ID research is located in a no man's land. In this line of thinking the appropriateness of different bibliometric indicators for measuring the impact of interdisciplinary research is being studied (Rinia et al., 2002) and recommendations to improve the assessment of ID efforts are being published by different policy advisory boards (Metzger and Zare, 1999; Grigg, 1999).

3. APPROACHES TO THE STUDY OF INTERDISCIPLINARITY

In the framework of bibliometrics different approaches to the study of interdisciplinary research have emerged during the last years. This can be seen through the review of the numerous and relevant studies on cross-disciplinary research that have been published in the scientific literature.

Most of the studies focus on the micro or meso level, macro level analyses being less frequent. The study of interdisciplinarity in publications can be approached from different perspectives such as the following: collaboration amongst authors from different disciplines, co-occurrence of several classification codes in publications, interdisciplinary nature of publication journals and cross-disciplinary references and citations.

3.1 Collaboration between Authors with Different Academic Training or Background

Interdisciplinary research does not necessarily imply collaboration between researchers from different disciplines, but this is, in fact, one of the main sources of interdisciplinarity. Collaboration amongst scientists from different disciplines is widespread, but measuring it is not easy.

The academic training of scientists cannot be determined by their publications, but it can be obtained by surveys or questionnaires. Collaboration across disciplines was studied in the survey conducted in 1984 among 5600 scientists from 10 different European countries (Franklin, 1988), in which 85% of the respondents stressed that the most promising research directions involved multi-disciplinary work. In a recent survey of approximately 600 Spanish research teams around 60% of Pharmacological and Cardiovascular teams and 45% of Materials Science described themselves as multidisciplinary, attending to the professional background of the group members (Sanz et al. 2001). It is interesting to note that more than 80% of the teams said that they used knowledge and techniques from other disciplines. Moreover, the journals they read and where they publish span a wide range of disciplinary categories and only a third of the journals used actually belonged to their main specialization field. Collaboration with other groups was present in more than 70% of the groups, although it was not described as interdisciplinary in all cases. The area with fewer interdisciplinary groups in terms of their composition was Materials Science, which nevertheless showed higher intensity of external research collaboration than the other areas. This was explained by the fact that single-disciplinary teams require external collaboration more frequently to gain access to information beyond their own discipline.

In many cases direct information provided by the authors through surveys or interviews is not available. However, we can rely on bibliometrics to measure collaboration amongst authors from departments or centres of different disciplines, as an indicator of interdisciplinary collaboration. This approach implies some limitations, since we do know that interdisciplinary collaboration is not always reflected in authorship and that the authors' organisational affiliations do not always represent their specialties. For example, two scientists working in the same department may have been trained in different disciplines and we are not able to detect this collaboration as interdisciplinary according to affiliation criteria. On the other hand, two scientists with the same training may be working in different disciplines according to their organizational affiliation. In spite of this limitation the method has proved useful for gaining insight into cross-disciplinary research through publication analyses (see for example Qiu, 1992; Hinze, 1999; Qin et al. 1997; Bordons et al., 1999).

The most frequently used indicator of cross-disciplinary collaboration is the *percentage of co-authored interdisciplinary papers*. In a study on the area of auto-immune diseases, the proportion of cross-disciplinary publications ranged from 43% to 66%, depending on the countries.

Moreover, more than 58% of all patents draw on scientific knowledge from more than just one discipline (Hinze, 1999). Amongst the limitations of the methodology, we can mention the difficulty of determining in advance which classification scheme of disciplines is going to be used. Very detailed classifications might overestimate interdisciplinary contacts, whilst the contrary applies for very general classifications. On the other hand, the percentage of co-authored interdisciplinary papers does not distinguish between a paper incorporating two disciplines and one incorporating four; however, this limitation could be overcome by introducing other indicators to measure the scope of interdisciplinarity.

3.2 Presence of Keywords or Classification Codes from Different Disciplines in Documents

The co-classification analysis has been used to describe the intellectual structure of research areas, with special emphasis on mutual relations between its subject fields. With this approach we can assume that ID is revealed by the presence of keywords and/or classification codes from different disciplines, so the degree of co-occurrence of subject classification headings will provide interesting information about interdisciplinarity in the area under analysis. A detailed presentation of the advantages and limitations of this methodology can be read in a study on energy research (Tijssen, 1992). According to this author the *rate of interdisciplinarity* in a given area can be obtained from the sum of all co-occurrence frequencies divided by the total of all occurrence frequencies. Moreover, we can measure how each field contributes to the total interdisciplinarity of a given area, or which is the rate of interdisciplinarity of each field. Finally, it enables us to analyse the relationship between fields, to visualize it graphically in maps, and to study its evolution over time. Co-classification relations can be represented graphically by means of different multivariate analyses, which result in the construction of 'maps' in which the fields appear as points in the map, linked by lines that represent the ID relations. This method can be useful for identifying emerging interdisciplinary specialties or subfields as well as to detect links between science and technology.

Other authors have used co-classification analyses in order to study the structure of different areas (see for example McCain, 1995). Amongst the limitations of the methodology we can mention that its validity depends on the adequate coverage of the database used, the frequent update of the classification system, and the expertise of the indexers.

3.3 Interdisciplinarity through Publication Journals

The distribution of the scientific output of a given centre by disciplines is useful for obtaining its thematic profile and enables us to know whether it is concentrated on a given discipline or whether it shows a stronger multidisciplinary activity. Different databases provide classification schemes of journals by disciplines. For example, the *Science Citation Index* classification includes more than 4000 journals into 150 categories. Thus, documents can be assigned to the discipline of their publication journal.

In this context *the percentage of documents published outside its main research area* by a given centre or a given group can be used as an indicator of interdisciplinarity, and *boundary crossing authors*, described as those who published in journals from disciplines outside their own, can be identified. In the area of Pharmacology it was observed that 45% of the publications of the Spanish pharmacologists were published in non-pharmacology journals (Bordons and Barrigon, 1992). However, this indicator shows large differences depending on the areas analysed. In a study on Sociology and Political Science, boundary crossing authors were found to come preferably from neighbouring disciplines and succeed in achieving interdisciplinary information transfer, as measured through their papers´ citation rates (Pierce, 1999). An interesting study on publications of the Australian University showed that almost 70% of Physical Sciences and Chemical Sciences departments published in journals belonging to their own fields, whilst only 37% of Mathematical Sciences and 50% of Biological or Agricultural Sciences were published in journals of their corresponding categories (Bourke and Butler, 1998). In fact, they found that researchers in academic departments publish in journals across a range of fields outside their nominal disciplinary affiliation: Mathematics departments publish in Mathematics and Physics, whereas Applied Sciences departments show a wide range of publication disciplines. Important relationships between Biology, Agriculture, and Medical Sciences were also found.

Several classifications of journals in categories, such as the ISI one, accept multi-assignation of journals, that is, journals can be classified into more than one category. The assignment of journals to categories is based on a review of the journals content, as well as on the analysis of emergent patterns in cited/citing journals: both an objective and a subjective criterion are used, since they complement each other. Those journals which are assigned to more than one category are to be read by different communities of scientists, so they must presumably include knowledge useful for different disciplines, that is, interdisciplinary knowledge. Under this assumption, several indicators based on multi-assignation of journals have been introduced for the study of interdisciplinarity.

ISI multi-assignation of journals into subject categories was proposed by Katz and Hicks (1995) as a measure of interdisciplinarity. They developed a journal classification scheme derived from the ISI and adapted to measure separately publication output in interdisciplinary journals. More recently, Morillo et al. have introduced a set of indicators based on multi-assignation of journals with the purpose of obtaining a general view of research areas and disciplines. The usefulness of ISI multi-assignation of journals as a measure of interdisciplinarity was analysed first in the area of Chemistry (Morillo et al., 2001) and later with a more global approach (Morillo et al., 2003).

3.4 Knowledge Transfer between Disciplines

Knowledge flows between disciplines can be analysed through different approaches. The movement of a scientist from one discipline to another is perhaps the most efficient way of transferring knowledge and that is why migration patterns amongst disciplines have been studied to gain insight into the relationships between disciplines. From the bibliometric point of view the distribution of references and citations amongst disciplines has proved useful for the study of information flows amongst disciplines. It is considered that a very influential discipline which provides concepts and techniques to another one will be heavily cited in the publications of the latter. So that, the field breakdown of all publications cited or referenced in a specific field will provide an interesting overview of its interdisciplinary profile. Moreover, this methodology applies to both science and technology. Citations to scientific papers in patents have been used successfully to quantify links between scientific and technological fields of knowledge (Narin et al., 1997).

3.4.1 Migration of scholars

The movement of scholars from one field to another is often accompanied by a flow of ideas in the same direction. Thus, the analysis of scholar migration patterns provides information about the relations between fields, since when two fields exchange large numbers of scholars it is probably because they share important cognitive features. Under these assumptions Hargens (1986) studied the migration of scholars between fields. As a result 17 general areas of scholarship are grouped into two clusters: the Natural Sciences and Mathematics, and the Behavioural Sciences and Humanities. Within the first cluster, Mathematics, Astronomy and Computer Science stay at one pole and Life Sciences at the other. Physics and Mathematics are central fields amongst the first group, and

Experimental Biology is a central field amongst the second group. Chemistry occupies an important structural position as an intermediate link between the Physical and Life Sciences, as was also found by means of citation flow analysis by other authors (Small and Griffith, 1974; Narin et al., 1972). Fields in the Behavioural Sciences and Humanities cluster show a strong trend of exchanging scholars amongst themselves rather than with fields in the first cluster. The migration patterns found suggest that the primary force shaping relations between scholarly disciplines is that of common cognitive issues: the degree to which fields are concerned with similar topics and investigate them with similar concepts and methods.

According to Hargens the information flows between disciplines are reciprocated, so hierarchical relationships are weak or absent between disciplines. However, other authors have found that the relationships between disciplines are asymmetric. It means that some fields are more fundamental than others, in the sense that they supply more information to other fields than they receive from them. Thus the description of fields as either 'donors' or 'receivers' becomes interesting, based on the prevailing direction of scientist's movements: sending scientists to other fields or receiving them. A study on Dutch universities (Le Pair, 1980) showed that some fields are predominantly donors (Pharmacy, Physics, Chemistry, Biology) whereas others are mainly receivers (Medicine, Agricultural Sciences).

A study of information flows amongst Social Sciences and Humanities disciplines in Japan (Urata, 1990) found that Philosophy, Psychology, History, and Linguistics were 'donor disciplines', as they offer a large amount of information to other disciplines, whereas Education and Sociology were 'receivers', since they obtain a great deal of information from other disciplines. The 'receiver' areas, also called 'importing areas' by Cronin (1990), are those less independent areas, since they rely heavily on other disciplines (they are the most interdisciplinary disciplines), whilst the 'donor' or 'exporting areas' are more robust and independent.

3.4.2 Distribution of references/citations over categories

The disciplinary distribution of the references of a given paper enables us to identify the main related disciplines in which the new knowledge is based. This approach can be applied to the study of individual papers, as shown by Glänzel et al. (1999) in the classification of papers published in multi-disciplinary and general journals using reference analysis. This methodology can also be applied to the journal level. The disciplinary profile of a journal can be obtained and compared with the corresponding profile of other journals or with the profile of the same journal in another time period.

Moreover, apart from obtaining the list of the most related disciplines, it is possible to quantify the interdisciplinary behaviour of a journal through different diversity indexes, such as the Pratt or the Brillouin indexes (see for example Morillo et al., 2001; Steele and Stier, 2000).

The most successful bibliometric indicator for the study of cross-disciplinary research is the *citations outside category* (COC), first introduced by Porter and Chubin (1985) in a case study on Demography, Operations Research/Management Science, and Toxicology. With this approach a citation is classified as COC when the subject matter of the cited journal is different from that of the citing journal. The usefulness of the indicator was validated by the judgment of peers who described the degree of ID assigning papers to one out of five categories of increasing interdisciplinary nature. As a result almost 70% of the citations belonged to the same discipline of the journal analysed, although inter-discipline differences were found, Toxicology being the most interdisciplinary. According to the results of Porter and Chubin this indicator, *citations outside category,* is quite robust within a category across journals and within journals over time. Moreover, citations between distant categories, that is, across broad field categories (Engineering, Life Sciences, Physical Sciences and Social Sciences) are extremely infrequent.

More recently an *interdisciplinarity index* for journals was proposed and applied in the area of Fertility by Tomov and Mutafov (1996). In this case citations and references outside category were considered together with the number of citing and cited journals. The index was calculated as "the sum of the ratios between the number of journals from all other disciplines and the number of journals from the same discipline cited by the journal or citing it, on the one hand, and of ratios between the numbers of citations to and by the abovementioned journals on the other". However, the authors stress that scientists, journal editors, and research policy managers are usually more interested in the identification of those fields which are more related to any given discipline than in obtaining an ID index.

The study of cross-disciplinary citations at the macro level of the global research system has been tackled in a few studies. In the Science and Engineering Indicators Report 2000 citations during 1997 in US scientific papers across 11 broad areas were analysed (NSB, 2000). It is interesting to note that only a minority of the references were 'external', that is, referred to another broad area (COA = *citations outside area*), showing values ranging from 17% in Physics and Earth & Space Sciences to 39% in Biology and Engineering. These results are consistent with those from an exploratory study of knowledge exchange between disciplines published recently by Rinia et al. (2002), who analysed world publication output in the 1999 SCI. Cross-disciplinary citations in journal articles were studied, considering 167 categories and 17 broad areas. In almost all disciplines the largest share of references was given to publications of their own discipline. With regard to broad areas, Physics showed the highest dependence on results published in literature from its own discipline, with a low knowledge import from other fields; and Biology and Engineering appeared as high receivers, as found in the NSB study above mentioned. Differences in the number of broad areas considered in both studies (11 areas in the NSB study and 17 in the Rinia one) hinder detailed comparisons, but it seems that journals in the basic Life Sciences were the most important source of external knowledge for other fields.

In relation to the citations outside category, Van Leeuwen and Tijssen (2000) have analysed citations given by research papers in 1985–1995, using the ISI classification of journals into 119 categories. On average 69% of the citations were cross-disciplinary, that is, referred to another category (COC). This value is quite high compared to the abovementioned rates at the aggregate level of broad areas, that is, a higher interdisciplinarity is detected at the level of categories. This can be explained by COC including citation links amongst close categories which belong to the same broad area. It is interesting to note that important differences were found amongst disciplines and that the majority of the top–ranking disciplines belonged to the biomedical sciences. Surprisingly, the percentage of cross-disciplinary citations did not change significantly during the ten years under analysis, in spite of the increasing interdisciplinary nature of science pointed out even by scientists themselves.

Since the size of the citing and cited discipline and the citation characteristics of the fields concerned may play a role in the final COC, new indicators have been recently introduced. That is the case of the *relative openness*, which measures the preference of a scientist in a discipline for results published in journals in other disciplines, and the *import/export ratio*, which is the number of external citations given to publications of a discipline divided by the number of publications of this discipline (Rinia et al., 2002).

4. AN OVERVIEW OF SCIENCE THROUGH THE ISI CLASSIFICATION OF JOURNALS INTO CATEGORIES

Results obtained in studies on interdisciplinarity at the micro or meso level cannot be generalised into other fields. That is why studies at the macro level, which provide an overview of all fields in science as well as the main relationships between them, are especially interesting. Reviewing the published literature, three main approaches have succeeded in obtaining this type of general overview: studies on mobility of scientists across fields, studies on cross-disciplinary journals, and analysis of cross-disciplinary citations. In this section we will focus on the usefulness of cross-disciplinary journals, defined as those that are multi-assigned in more than one category, for the study of science structure.

A tentative typology of disciplines and research areas according to their degree of interdisciplinarity was presented recently (Morillo et al. 2003) considering several indicators based on *ISI multi-assignation of journals in subject categories*. Research areas and categories were described according to the quantity of their links (number of related categories) and their quality (diversity and strength of links). High levels of inter-relations between categories were observed, which was consistent with scientists' perceptions since discipline boundaries are always artificial and a high relationship does exist between knowledge in the different disciplines. *Multi-assignation rate* and *pattern* of research areas are shown in Table 19.1.

Table 19.1. Multi-assignation rate and pattern of research areas

Area	*No. Journals*	*% Mult Journals.*	*Multi-assignation pattern*		
			% Internal	*% External*	*% Int.& Ext.*
Engineering/Technology	1330	56.5	42.4	45.9	11.7
Biomedicine	1233	56.4	33.6	56.6	9.8
Physics	606	48.0	20.6	69.8	9.6
Chemistry	430	43.5	11.2	81.8	6.9
Agric./Biol./Environment	1085	43.3	33.8	57.7	8.5
Clinical Medicine	1461	41.2	30.9	60.6	8.5
Social Sciences	1564	39.2	56.6	35.7	7.7
Mathematics	337	38.9	19.8	71.8	8.4
Arts/Humanities	1192	11.0	44.3	54.2	1.5

According to multi-assignation percentage Arts & Humanities was the most isolated area (only 11% of its journals are multi-assigned), whereas Engineering and Biomedicine were in the opposite situation (multi-assignation rate of 56%). Concerning Engineering, 46% of its multi-assigned journals were shared with categories from other areas different from Engineering (external links) while 42% were shared with more close categories within the Engineering area (internal links). Around 12% of the journals were included in Engineering categories and in at least one other category from another area. Some outstanding results are the low multi-assignation rate of Humanities and the high internal multi-assignation of Social Sciences, whose categories establish links principally with other categories of their own area.

It is surprising to find both a basic and an applied area are the ones showing the highest interdisciplinarity. This is better explained when comparing our results with those obtained using other indicators. In a previous study (Rinia et al., 2002) the area with the highest interdisciplinarity as measured through citations received corresponded to Biomedicine, a basic area, whilst Engineering/Technology, an applied area, was the first when considering references given to other areas. That is, Biomedicine is acting as a 'donor' exporting knowledge, and Engineering is a 'receiver' of knowledge. These results are consistent with ours based on multi-assignation, as we are considering both sort of links between disciplines, not differentiating the direction of knowledge flows. To quantify the diversity of relationships the number of different links established between pairs of categories was calculated. Since the larger the category size the larger is the diversity of relationships, the number of links per category size was calculated. The highest diversity of links corresponded to Engineering/Technology, which showed one link per two journals, and the lowest one to Arts/Humanities, with one link per eight journals.

The strength of links between areas, calculated through the Salton index, is shown in Figure 19.1. Areas are located in the figure according to their multi-assignation rate (vertical axis) and to their basic/applied nature (horizontal axis), measured through the average research level of the publication journals (Noma, 1986). Thus the highest multi-assignation rate corresponded to Biomedicine and Engineering, at the top of the figure, and the lowest one to Mathematics, at the bottom. The most important relations between areas are shown by lines: solid lines representing higher relationship than dotted lines. Considering shared journals between categories, the highest relationship was found between Biomedicine and Clinical Medicine, followed by that between Physics and Engineering/ Technology. Other related areas were Agriculture and Biomedicine, followed by Chemistry and Engineering/Technology. Circles in Figure 19.1 show the grouping of areas according to their multi-assignation percentage and pattern through hierarchical clustering analysis. Four different clusters of areas are identified. It is interesting to note that Arts/Humanities and Social Sciences remain in two separate clusters owing to their different behavioural pattern (they are not shown in the figure as the indicator basic/applied level is absent in most of their journals).

Categories were grouped into clusters according to their *multi-assignation percentage, percentage of external links, and diversity of links*. Four different types of categories were found with a decreasing interdisciplinary nature:

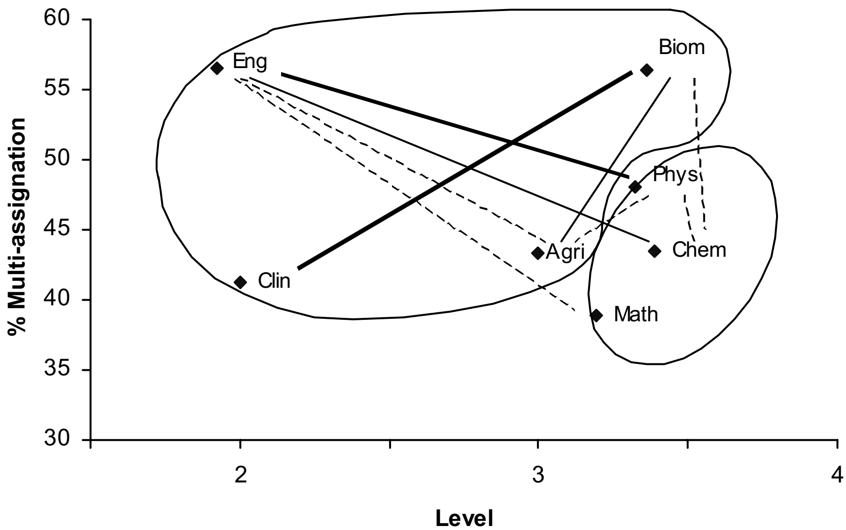


Figure 19.1. Strength of links between research areas
Agri= Agriculture, Biology and Environment; Biom= Biomedicine; Chem= Chemistry; Clin= Clinical Medicine; Eng= Engineering/Technology; Math= Mathematics; Phys= Physics.

- Type a. Categories with 'big interdisciplinarity' (25% of the total): this group of categories show high multi-assignation, especially the external type. Distant categories are linked together and relationships across areas are established. Categories such as Thermodynamics, Biological Psychology, Mathematical Psychology, Biotechnology or Environmental Sciences are included in this group.
- Type b. Categories with 'small interdisciplinarity'[1] (33% of the total): high multi-assignation and predominantly internal relationships. In this cluster categories are mainly related to other categories within their own area. This is the case of Instrumentation, which shows links with other categories within Engineering, or Limnology, which is related to other categories in Agriculture.
- Type c. Categories with mid low interdisciplinarity (25%): low multi-assignation percentages, but with a predominance of external links. Information and Library Science category is located in this cluster.
- Type d. Categories with low interdisciplinarity (19%): this group includes categories with low multi-assignation rate and mostly internal

[1] The terminology of 'small' and 'big' ID is taken from Schmoch et al. (1994).

links. It should be noted that 70% of the Humanities categories belong to this group.

The behaviour of some particular disciplines is shown in Table 19.2. Environmental Sciences and Biotechnology appear as highly interdisciplinary categories (type a), with a high percentage of journals multi-assigned to more than one discipline, and high diversity and strength of links. Regarding Environmental Sciences, it is interesting to point out that it shares journals with 45 different categories, distributed through all the different areas. According to the number of shared journals, it is mainly related to Engineering Environment, Toxicology, Water Resources, Ecology, and Public Health. Considering strength of links, the closest categories are the five mentioned above together with Limnology and Meteorology (two small categories whose importance is not clearly seen through the absolute number of shared journals). Biotechnology is mainly related to Biochemistry & Molecular Biology, Food Science & Technology, Genetics and Microbiology, measured either by number of shared journals or by strength of links.

Table 19.2. Description of some categories through ID indicators

Category	*Cat. Type*	*No. Journals*	*No. Related categ.*	*No. Related areas*	*% Multi-assign.*	*Diversity of links*	*Strength of links*	*Multi-assign. % Int.*	*%Ext.*	*Pattern % Int. Ext*
Thermodynamics	a	28	10	3	100%	0.357	9.244	4%	89%	7%
Envir.Sciences	a	120	45	9	77%	0.375	4.195	17%	50%	6%
Biotechnology	a	108	29	5	74%	0.268	4.084	18%	47%	13%
Transplantation	b	9	9	2	100%	1.000	6.712	56%	11%	33%
Instrumentation	b	45	27	6	96%	0.600	4.488	58%	27%	11%
Inf.Librar.Science	c	60	12	4	43%	0.233	3.800	19%	73%	8%
Literature	d	131	5	1	11%	0.050	2.200	100%	0%	0%

% Multi-assign. = % Multi-assignation of journals.

Concerning highly interdisciplinary categories in which internal links predominate (type b), we can see the categories of Transplantation and Instrumentation as an example. Transplantation is mainly related to other Clinical Medicine categories, such as Surgery and Nephrology. Instrumentation is a Physics discipline, related mainly to Physics and Engineering categories.

Information and Library Science is shown as an example of mid low interdisciplinary category (type c). It is mostly linked with Computer Information Systems, according to both number of shared journals and strength of links. This category displays links with 12 other categories from

4 different areas (6 categories in Social Sciences, 3 in Engineering, 2 in Arts/Humanities, and 1 in Clinical Medicine).

Finally, Literature is a sample of a low ID category (type d), with a low multi-assignation rate, low strength and diversity of links, and an internal pattern of multi-assignation. It is mainly related to History and Philosophy.

In summary, it is possible to describe disciplines through indicators related to multi-assignation, to obtain a general picture of their behaviour and to identify the main related categories. Moreover, differences over the years can be analysed. It is interesting to remark that the new categories added to the SCI classification over a 15–year period were significantly more interdisciplinary, according to multi-assignation indicators, than the rest of the categories. In fact, 80% of the new categories identified from 1981 to 1996 were located in highly interdisciplinary groups of disciplines (types a and b). A large number of these new disciplines are related to Technology: Biotechnology, several Materials Sciences and Engineering categories. Perhaps Nanotechnology will be a new category in the near future and expectedly with high ID. This technology is applied to very different areas and contributes to the blurring of borders between Physics/Chemistry and between Science/Technology, making synergistic interactions between scientists possible.

When comparing the findings from migration patterns, citation flows and multi-assignation patterns analyses, the results are not always consistent. However, the fact that different subject classification schemes are used in the published studies make comparisons difficult. Similar results were obtained by citation analysis and migration patterns analysis in a study of the Social Sciences and Humanities in Japan (Urata, 1990). The analyses of scholars' migration patterns proved especially useful for the study of large fields and less suitable for small ones, since reliable data cannot be obtained in the latter owing to the scarcity of scholars. For the study of small fields analyses based on citation flow such as co-citations seem more suitable. Thus data on citation flow and migration are complementary in that they are suited to analyses on different levels of scientific activity.

Similarly, the ID results obtained through the ISI multi-assignation indicators were compared to those obtained from a specialised database in the area of Chemistry (Morillo et al., 2001). Convergence between ISI multi-assignation rate of journals into categories, extra-disciplinary citations and multi-assignation of documents into Chemical Abstract sections was found in that study. The most appropriate indicators differ according to the level of analysis: areas, disciplines or journals. The most useful indicators at the area and specific discipline level were those based on ISI multi-assignation. In the case of ISI general or horizontal categories (such as Applied Chemistry), or at the level of journals, the most sensitive and appropriate indicators were those based on citing/cited patterns or those based on a specialised database classification. A paper by paper assignment was the most precise measure at the micro level, particularly for the study of journals: the study of single or multiple section assignment of papers gave as a result a typology of the analysed journals, which were described as specialised, multidisciplinary or interdisciplinary journals (Morillo et al., 2001).

5. CONCLUSIONS

At present, studies on interdisciplinarity from all possible perspectives are being increasingly requested. The study of the interdisciplinary process 'per se' has become a necessity and there is a growing body of literature dealing with its main opportunities and problems (Klein, 1996; Grigg, 1999; Weingart and Stehr, 1999). Cognitive and institutional obstacles for the development of interdisciplinary research are identified and studied. Science policy efforts focus on: establishing adequate criteria for the assessment of cross-disciplinary research; methods of strengthening and fostering interdisciplinary collaboration; and the identification of requirements needed for successful cross-disciplinary research.

Bibliometric approaches have proved useful for providing insight into specific aspects of cross-disciplinary research and for complementing other methodologies. However, the development of bibliometric studies is not devoid of problems. The lack of consensus about what should be considered as interdisciplinarity or the diversity of classification schemes used in the different studies are some of the problems to be solved. Moreover, the importance of maintaining up to date thematic classifications, such as the classification of journals into categories, is evident, since they are the basis for many of the bibliometric indicators designed to measure cross-disciplinarity.

We know that there is an increasing mutual dependence amongst disciplines in science that requires a good flow of knowledge beyond disciplinary boundaries. Disciplines are losing their identity and frontiers between science and technology are fading. By transcending the cognitive boundaries of existing disciplines, and looking at a scientific or technological problem from several different points of view, important scientific advancements can be achieved. A balanced development of all disciplines in science and technology is needed, not only for the advancement of established fields, but also for enhancing the unexpected emergence of new interdisciplinary fields (Song, 2003). Our results suggest that Engineering/Technology plays an important role in ID research: a high number of SCI new categories are classified under this area and their categories show a high number of relations with close and distant disciplines.

Instruments used in the study and evaluation of science should advance and evolve to reflect adequately the structure of modern science. Difficulties in the evaluation of interdisciplinary research by means of conventional methods have been pointed out by different authors. The creation of new and better bibliometric indicators for undertaking fair and rigorous assessment of this type of research and to obtain in depth studies of cross-disciplinarity in science is still a challenge to be taken up in the near future.

ACKNOWLEDGEMENTS

Some of the results presented here were financed by the research project SEC97–1375 and the research grant FPI–95 supported by the Spanish Ministry of Education and Culture.

REFERENCES

Bordons, M., Barrigón, S. (1992). Bibliometric analysis of publications of Spanish pharmacologists in the SCI (1984–89). II. Contribution to subfields other than ‘Pharmacology & Pharmacy’ (ISI). *Scientometrics*, 25, 425–446.

Bordons, M., Zulueta, M.A., Romero, F., Barrigón, S. (1999). Measuring interdisciplinary collaboration within a University: the effects of the Multidisciplinary Research Programme. *Scientometrics*, 46, 383–398.

Bourke, P., Butler, L. (1998). Institutions and the map of science: matching University departments and fields of research. *Research Policy*, 26, 711–718.

Braun, T., Schubert, A. (2003) A quantitative view on the coming of age of interdisciplinarity in the sciences 1980–1999. *Scientometrics*, 58, 183–189.

Cronin, B., Person, S. (1990). The export of ideas from information science. *Journal of Information Science*, 16, 381–391.

Franklin, M.N. (1988). *The community of science in Europe*. Brussels: Commission of the European Communities.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M. (1994). *The new production of knowledge*. London: Sage.

Glänzel, W., Schubert, A., Czerwon, H.J. (1999). An item–by–item subject classification of papers published in multidisciplinary and general journals using reference analysis. *Scientometrics*, 44, 427–439.

Grigg, L. (1999). Cross-disciplinary research. A discussion paper. Commissioned Report No.61. Canberra: Australian Research Council.

Hargens, L.L. (1986). Migration patterns of U.S. Ph.D.s among disciplines and Specialties. *Scientometrics*, 9, 145–164.

Hinze, S. (1999). Collaboration and cross-disciplinarity in autoimmune diseases. *Scientometrics*, 46, 457–471.

Katz, S., Hicks, D. (1995). *The classification of interdisciplinary journals: a new approach.* In M.E.D. Koenig, A. Bookstein (Eds.), Proceedings of the Fifth Biennial Conference of the International Society for Scientometrics and Informetrics (pp. 245–254). Medford: Learned Information.

Klein, J.T. (1996). Interdisciplinary needs: the current context. *Library Trends*, 45, 134–154.

Le Pair, C. (1980). Switching between academic disciplines in universities in the Netherlands. *Scientometrics*, 2, 177–191.

McCain, K.W. (1995). The structure of biotechnology R&D. *Scientometrics*, 32, 153–175.

Metzger, N., Zare, R.N. (1999). Interdisciplinary research: from belief to reality. *Science*, 283, 642–643.

Morillo, F., Bordons, M., Gómez, I. (2001). An approach to interdisciplinarity through bibliometric indicators. *Scientometrics*, 51, 203–222.

Morillo, F., Bordons, M., Gómez, I. (2003). Interdisciplinarity in science: a tentative typology of disciplines and research areas. *Journal of the American Society for Information Science and Technology*, 54, 1237–1249.

National Science Board (2000). Science and Engineering Indicators 2000. Arlington, VA: National Science Foundation, (NSB–00–1).

Narin, F., Carpenter, M., Berlt, N.C. (1972). Interrelationships of scientific journals. *Journal of the American Society for Information Science*, 23, 323–331.

Narin, F., Hamilton, K.S., Olivastro, D. (1997). The increasing linkage between U.S. technology and public science. *Research Policy*, 26, 317–330.

Noma, E. (1986). *Subject classification and influence weights for 3,000 journals*. Report under Contract No. NIH-N01-OD-5-2118. New Jersey: CHI Research, Inc. (updated in 1999).

OECD (1998). *Interdisciplinarity in Science and Technology*. Directorate for Science, Technology and Industry. Paris: OECD.

Pierce, S.J. (1999). Boundary crossing in research literatures as a measure of interdisciplinary information transfer. *Journal of the American Society for Information Science*, 50, 271–279.

Porter, A.L., Chubin, D.E. (1985). An indicator of cross-disciplinary research. *Scientometrics*, 8, 161–176.

Qin, J., Lancaster, F.W., Allen, B. (1997). Types and levels of collaboration in interdisciplinary research in the sciences. *Journal of the American Society for Information Science*, 48, 893–916.

Qiu, L. (1992). A study of interdisciplinary research collaboration. *Research Evaluation*, 2, 169–175.

Rinia, E.J., Van Leeuwen, T.N., Bruins, E.E.W., Van Vuren, H.G., Van Raan, A.F.J. (2002). Measuring knowledge transfer between fields of science. *Scientometrics*, 54, 347–362.

Sanz, L., Bordons, M., Zulueta, M.A. (2001). Interdisciplinarity as a multidimensional concept: its measure in three different research areas. *Research Evaluation*, 10, 47–58.

Schmoch, U., Breiner, S., Cuhls, K., Hinze, S., Münt, G. (1994). *Interdisciplinary cooperation of research teams in science–intensive areas of technology*. Final report to the Commission of the European Unit (VALUE II, Interface II, HS1). Karlsruhe: Fraunhofer Institute for Systems and Innovation Research.

Small, H. (1999). A passage through science: crossing disciplinary boundaries. *Library Trends*, 48, 72–108.

Small, H., Griffith, B.C. (1974). The structure of scientific literatures. I. Identifying and graphing specialties. *Science Studies*, 4, 17–40.

Song, C.H. (2003). Interdisciplinarity ad knowledge inflow/outflow structure among science and engineering research in Korea. *Scientometrics*, 58, 129–149.

Steele, T.W., Stier, J.C. (2000). The impact of interdisciplinary research in the environmental sciences: a forestry case study. *Journal of the American Society for Information Science*, 51, 476–484.

Tijssen, R.J.W. (1992). A quantitative assessment of interdisciplinary structures in science and technology: co-classification analysis of energy research. *Research Policy*, 22, 27–44.

Tomov, D.T., Mutafov, H.G. (1996). Comparative indicators of interdisciplinary in modern science. *Scientometrics*, 37, 267–278.

Urata, H. (1990). Information flows among academic disciplines in Japan. *Scientometrics*, 18, 309–319.

Van der Besselaar, P., Heimeriks, G. (2001). *Disciplinary, multidisciplinary, interdisciplinary. Concepts and indicators*. In M. Davis and C.S. Wilson (Eds.), Proceedings of the 8th International Conference on Scientometrics and Informetrics (pp. 705–716). Sydney: University of New South Wales.

Van Leeuwen, T., Tijssen, R. (2000). Interdisciplinary dynamics of modern science: analysis of cross-disciplinary citation flows. *Research Evaluation*, 9, 183–187.

Weingart, P., Stehr, N. (Eds.). (1999). *Practising interdisciplinarity*. Toronto: University of Toronto Press.

Chapter 20

CITATIONS TO PAPERS FROM OTHER DOCUMENTS

Evaluation of the Practical Effects of Biomedical Research

Grant Lewison

Department of Information Science, City University, London, EC1V 0HB, England
E-mail: g.lewison@soi.city.ac.uk

Abstract: Citations to biomedical research papers from different types of document — clinical guidelines, textbooks, government policy documents, international or national regulations and newspaper articles — can provide new indicators of the utility of such research. However, most such citing documents will be national in character, and in order to provide an international perspective it will be necessary to combine several databases constructed to the same protocols and linked through the Web.

1. INTRODUCTION

Conventional bibliometric analysis has, up until quite recently, focussed on the extent to which individual scientific papers, or groups of them, are cited by other papers in the serial literature. The creation of the *Science Citation Index* (SCI) by Eugene Garfield in 1962, initially focussed just on genetics (Garfield, 1955, 1979; Lederberg, 2000; Thackray and Brock, 2000), has enabled the determination of citation counts to be conducted readily, and a vast literature has developed. Most of this is concerned with the tabulation of citation count numbers (van Raan, 2000; Ingwersen et al., 2000) and various ratios (Braun and Glänzel, 2000; Moed, 2000; Shama et al., 2000; Glänzel, 2000), which purport to allow the comparative merits of different groups of papers to be observed. The literature on citation theory is somewhat more restricted; it began with Garfield (1963) and modern commentators include Peritz (1992), Kostoff, Leydesdorff, van Raan, Vinkler (1998), and Wouters (1999), and indeed a recent study (Hanney et al., 2003) has cast doubt on whether cited papers are of major importance to

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 457-472.

the citing works, and on whether citation counts correlate with the long-term practical influence of individual biomedical papers.

While it would be unkind to compare the use of citations for research evaluation to the search for his keys by a drunken man under a street light, it has to be admitted that the ready availability of the SCI has tended to limit the search for other evidence of the impact of research. An exception has to be made for the study of patents. The analysis of non-patent references (NPRs) on US and European patents has now been organised with commendable zeal by CHI Research Inc. in the USA (Narin, 1994; Harhoff et al., 1999) and by some European research groups, e.g., those at Leiden (Noyons et al., 1998; Tijssen, 2001) and Karlsruhe (Schmoch, 1993; Grupp and Schmoch, 1999). Patent references have the big advantage that they are carefully considered by the applicants and the examiners because they both underpin and limit the claims being made for novelty. On the other hand, rather few scientific papers, even ones in basic research and in advanced sub-fields such as human genetics (Anderson et al., 1996), are ever cited by a patent. This contrasts with the situation with citation by papers. Most biomedical papers receive at least a handful of citations in their first five years after publication (Lewison, 2003), so furnishing a convenient yardstick that can be applied to almost all evaluation exercises.

Nevertheless, both paper and patent citation counts may be inappropriate as measures of the practical effects of biomedical research. In the first place, they are only distantly related to the provision of better health care. Figure 20.1 shows a possible model for how research can impact on healthcare. It is inevitably complex, and even this diagram omits some possible pathways, and factors such as the provision of resources and their equitable distribution. The point is that better health depends on a wide variety of possible interventions, or reactions by the public that lead to a healthier lifestyle. Government policy, in turn influenced by public opinion and the news media, also plays a major role. Secondly, these two citation indicators greatly favour basic research over clinical work (Narin and Olivastro, 1998) and, *a fortiori*, socio-economic investigations, which have become relatively under-valued as a result (Chalmers and Sinclair, 1985). And thirdly, there are some built-in national biases in citation indicators that put a premium on publications in US journals and on papers co-authored with the USA (Lewison, 2003). This last criticism means that much research conducted outwith the USA will be evaluated according to its effect on US researchers rather than to people in the country where it was performed, who may well have been the intended audience, particularly in the more clinical or sociological subjects.

In this chapter we shall turn to some alternative indicators of research impact. They also depend on the linking of two documents through citation and, as with patent citations, they may establish the practical utility of only a small minority of biomedical papers. But they are, like patent citations, the products of processes of careful consideration, sometimes by an expert committee, so that some of these new indicators will share the often celebrated virtues of peer review (Moxham and Anderson, 1992; Wood, 1997).

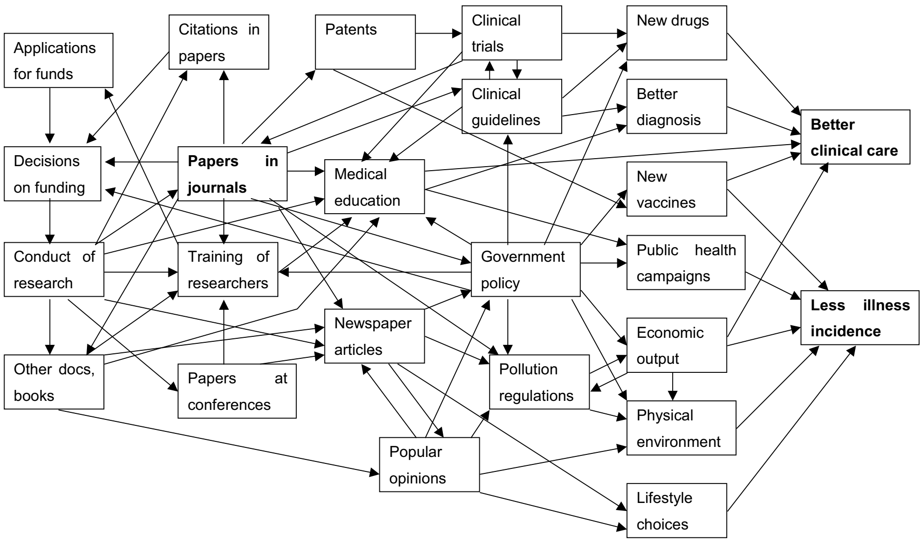


Figure 20.1. The links between research and better health

Where these new indicators differ is that they are, for the most part, national and will need to be determined separately in each country, though there may be some read–across to other nations. Whereas the SCI is a single commercial enterprise, these new indicators will need to be produced by a multi-national collaborative effort. Fortunately, the practical possibilities of remote data entry and access provided by the World Wide Web make this a feasible task, though we need to provide standard protocols and thesauruses for the purpose. There are also some important and difficult issues concerning intellectual property rights and finance that need to be resolved.

Briefly, we consider the listing and analysis of research publications cited on the following types of document.

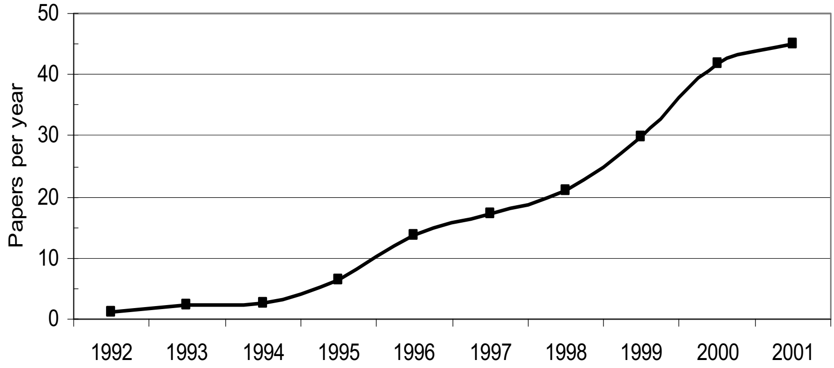


Figure 20.2. Numbers of SCI papers with 'evidence' and 'guideline' in their titles, three–year moving average

This is by no means a complete list, but it will suffice to demonstrate the methodology and bring out the problems entailed in the new activity. They are:

1. Clinical guidelines;
2. Medical textbooks;
3. Government policy documents;
4. International, regional and national regulations;
5. National and regional newspaper articles.

We will discuss each of these in turn, although the 'state of the art' differs greatly between them. Nevertheless, there are some methodological lessons that have been learned and which may usefully inform practice in the newer areas.

2. CLINICAL GUIDELINES

These documents are increasingly being developed and published, and (one hopes) used to inform and improve medical diagnosis and treatment. For example, Figure 20.2 shows the numbers of SCI papers with 'guideline(s)' and 'evidence' in their titles, which have risen rapidly since the mid-1990s. Many of these documents are, in effect, reviews of selected randomised and controlled clinical trials (RCTs), which are the 'Gold Standard' by which proposed new developments in clinical practice should be evaluated. However due allowance should be made for publication bias against those trials yielding negative results (Sterne et al., 2001; Schluchter, 2003).

Early work on the papers cited by a sample of UK clinical guidelines was carried out by Grant (1999, 2000). He showed that they depended almost entirely on clinical work (as opposed to basic research) and that British research was disproportionately cited in relation to its presence in the biomedical literature (about 10% on an integer count basis). There are, in fact several series of guidelines now current in the UK, including the recommendations of the Cochrane Collaboration (Dean, 2002; Laupacis, 2002) and a bi-annual book, *Clinical Evidence*, published by the BMJ.

There are also guidelines published by the Scottish Intercollegiate Guidelines Network, SIGN, which was set up in 1993 (Petrie et al., 1995; Petrie and Harlen, 1997) and has just issued No. 74, although some of the early ones have been withdrawn. In 1999 the National Institute for Clinical Excellence, NICE, was established (Dean, 1999) to produce guidelines to govern which treatments are deemed to be cost effective for use by the National Health Service (NHS). These are each based on a detailed Health Technology Assessment (HTA), nearly all carried out by one of five specialised units in English universities. The SIGN and NICE guidelines cover a large range of diseases and disorders, though most (60%) of the NICE guidelines are concerned with new drugs and many of these (nearly two fifths of them) are for cancer.

The references on the SIGN guidelines and NICE HTAs have all been processed by City University (Lewison and Wilcox–Jay, 2003) and matched to papers recorded in the SCI (or, occasionally, the Social Sciences Citation Index (SSCI)) whose bibliographic details have been retrieved and downloaded for analysis. This process is described below. The main purpose of this work, which is inevitably rather labour intensive in its present form, has been to identify those papers in the UK Research Outputs Database, ROD (Dawson et al., 1998) that have been cited on clinical guidelines. Details can then be given to members of the ROD club who pay to receive annual lists of papers acknowledging their support.

The current procedure for the recording of guideline references is as follows. First, the complete guideline is found on the Web and displayed as a pdf file. The title page and pages of references are printed. A cover sheet is then prepared with the years from 1980 (the first year of the SCI on CD-ROM) to the present listed on the left, and for each, spaces in which the reference numbers can be entered that correspond to journal publications (typically two thirds of them). When this sheet is complete, the papers are searched for on the SCI CD-ROMs, beginning with the earliest year, with a match being sought through three or four title words (with allowance for US spelling where necessary). Papers so found are 'collected' and their details are downloaded to a comma–delimited file, named 'SIGNxx' or 'NICEyy' as appropriate. Each file in turn is opened as an MS Excel spreadsheet, and an additional column added with the guideline number. The cited references are then combined into a single file, and this is further annotated with the publication year of each guideline. The source can be parsed to yield the journal name and the publication year of the cited reference. The file can then be analysed by standard means, in particular, geographically. UK papers are matched to those in the ROD in order to determine their funding sources (which have been previously individually determined by inspection in libraries).

The references turn out to be rather recent, with a median period between publication and citation on a guideline or HTA of 5.0 and 3.8 years respectively. They are overwhelmingly from the field of clinical medicine (94%), and clinical rather than basic as shown in Figure 20.3.

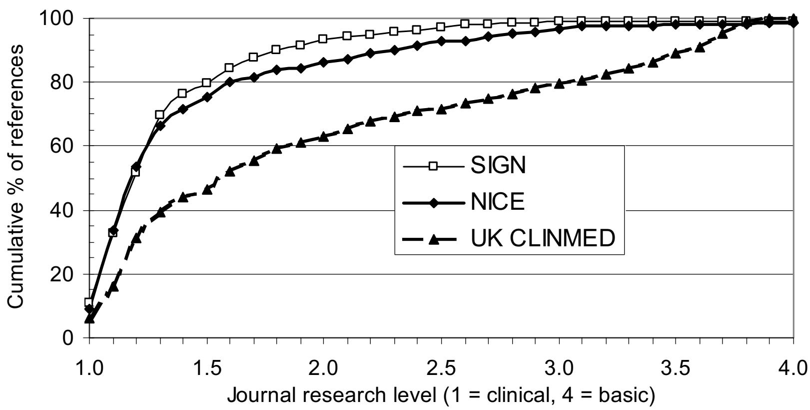


Figure 20.3. Cumulative Research Level distribution of papers cited on SIGN clinical guidelines (5–74) and NICE HTAs (1–58), and, for comparison, all UK clinical medicine papers 1995.

This shows their cumulative distributions by the research levels of the journals in which they have been published on a scale of 1 = clinical to 4 = basic research (Lewison and Paraje, 2003) with, for comparison, that of UK clinical medicine papers in the SCI in 1995 (the mean year for the references).

They are also in journals of relatively high citation impact, as determined by conventional means, particularly when due allowance is made for the journals being primarily clinical rather than basic

Although the references on these UK clinical guidelines are disproportionately taken from UK biomedical papers (in comparison with the UK presence in this literature of about 10%), the work of some other countries, particularly those in northern Europe, is also relatively over-cited. Figure 20.4 shows the contributions (integer counts) of some leading countries to these reference lists relative to their percentage presence in world biomedical literature from 1991–2000. [NO = Norway, FI = Finland, DK = Denmark, IE = Ireland, SE = Sweden, CA = Canada, AU = Australia, BE = Belgium, NL= Netherlands] It may therefore be possible to use this database of cited papers as a partial indicator of the relative utility of the clinical work from, say, Scandinavian medical schools. However, it would certainly need to be complemented with other, similar, data sets, based on clinical guidelines developed and used in those countries.

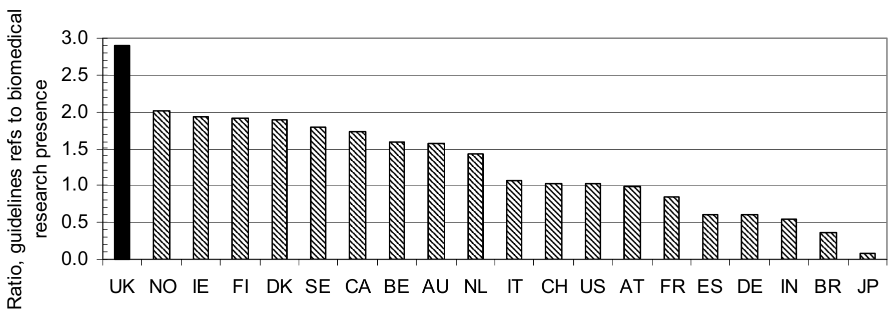


Figure 20 4. Percentage presence (integer counts) of 20 countries on SCI/SSCI papers cited by UK clinical guidelines and HTAs and in the biomedical research literature, 1991–2000.

3. MEDICAL TEXTBOOKS

The role of these in medical education is sometimes disputed (Rabow and McPhee, 2002) but it seems clear that they play an important role at least as references for students. Most of them nowadays consist of a series of individual chapters each written by an author selected by the editor(s) as being an expert in their sub-specialty. The chapters are furnished with references, which in principle can be listed on cover sheets, identified in the SCI or SSCI if they are journals covered by those databases, and their bibliographic details downloaded to file for analysis. Textbooks in specialised subjects often take the form of a collection of reviews so as to present a survey of the subject, and they have authors from many countries, chosen for their individual expertise, so their references may be expected to be similarly international.

One of the characteristics of textbooks is that they have to be kept up to date, and so revised editions are published from time to time, with some chapters extensively revised, re-written, or newly added. Consequently the lists of references will change with time as the state of the art in a particular subject advances, and some papers will no longer be cited. This gives the concept of citations a somewhat ephemeral quality, as a paper may be cited in one edition of a textbook but not in the next. Perhaps textbooks should be treated as if they were different volumes of a journal, coming out irregularly, so that a paper might be cited in some volumes but not in ones before or after. There might, in fact, be a 'citation window' during which certain scientific studies were regarded as providing the definitive judgement on a subject, but subject to being replaced by later ones, as happens with citations in journals.

Textbooks vary in the 'hardness' of the subject they cover, from those specialised in a particular disease such as schistosomiasis (Mahmoud, 2001) to those endeavouring to put medicine in a social context (Annandale and Hunt, 2000). References in the former book are almost exclusively (94%) taken from the serial literature; in the latter, books and other monographs dominate, and journal articles account for fewer than 40% of the references. The median age of the references is about 7 years; this applies both to books and journal articles in the latter textbook, see Figure 20.5.

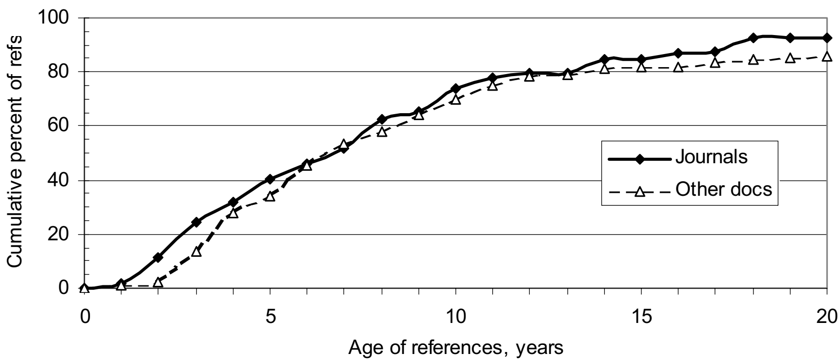


Figure 20.5. Ages of references in chapters 1–4 of Gender inequalities in health (2000)

Some textbooks with many references only give them in abbreviated format, without the title of the cited article and sometimes listing only the first author. This would make it slightly more difficult to identify the paper in the SCI or SSCI, as it would be necessary to match on both the author and the journal name, rather than just on title words. Moreover, the journal name is often abbreviated to a set of initials with which the reader is expected to be familiar.

4. GOVERNMENT POLICY DOCUMENTS

In view of the central role of government policy in the delivery of health care, shown in Figure 20.1, it is perhaps surprising that most such documents do not refer to an extensive evidence base underpinning policy. But there are some exceptions. In the UK the Department for Environment, Food and Rural Affairs (DEFRA) publishes a number of analyses of the effects of pollution, sometimes commissioned from universities, sometimes produced by expert working parties. These documents are all available for study on the Web, and many contain lists of references.

Although the large majority of these tend to be 'grey' literature — other reports and other government policy documents — there are also some research articles published in the peer–reviewed serial literature.

For example, the report, *Valuation of Air Pollution Effects on Ecosystems — a Scoping study*, prepared by the University of Aberdeen in September 2001, contained 82 references, of which 41 were journal articles. Of course, much of the cited literature is concerned more with the economic effects of pollution than with the health effects, but two are inter-linked. There is therefore a route by which research publications can be seen to influence government policy.

5. INTERNATIONAL, REGIONAL, AND NATIONAL REGULATIONS

Much industrial production and commercial business nowadays is regulated, principally for the protection of consumers and the environment. For example, there are international agreements on ionising radiation, with standards on acceptable doses published by the International Commission on Radiation Protection, a private non-profit body, founded in 1928 by the International Society of Radiology, registered in the UK and with a small secretariat in Sweden (http://www.icrp.org/). Its recommendations come with lists of references, mostly rather recent as shown in Figure 20.6: the median age of journal articles is 4 years and that of other cited documents, 6 years.

Another example is pesticide residues in food, which from time to time arouse public concern as possible sources of illness, particularly of cancer, or even survival (Colborn et al., 1996), although their true overall contribution to ill–health seems likely to be almost vanishingly small (Doll and Peto, 1981; McGinnis and Foege, 1993). Standards for these are based on 'evaluations' carried out by expert committees of the Codex Alimentarius Commission (CAC).

This Commission was formed jointly by the Food and Agriculture Organization (FAO) and the World Health Organization (WHO) in 1962 in order to promote international trade in foodstuffs.

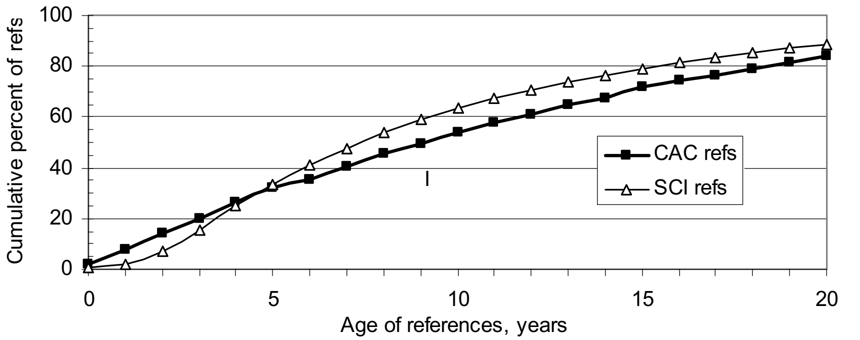


Figure 20.6. Dates of publications referenced in two draft ICRP Technical Guides, 2003

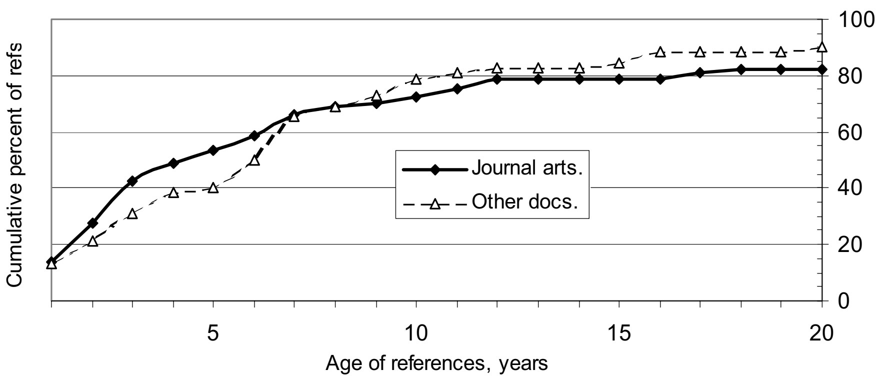


Figure 20.7. Ages of references in CAC pesticide evaluations (CAC) and in SCI pesticide papers (SCI)

Atiogbe (2001) examined a sample of the evaluations produced by the CAC on 18 frequently examined pesticides and the references in them in order to see which countries' research is used to develop these standards, how long it takes for papers to be cited, and what type of research is used. These references were compared with two further sets of papers, one taken directly from the SCI, selected by means of a pesticide 'filter' and published in 1991–98, and the other consisting of the references from the 1997 papers in the first set. Figure 20.7 shows that these papers are more recent than the references cited in the CAC pesticide evaluations, showing that the latter are not influencing practice until some years after they are published (median age, 9 years compared with 7 years for SCI pesticide citations). She also found that there was a good correlation between national contributions to the research cited on the CAC evaluations and their presence in the pesticide literature, even for some developing countries with relatively little scientific output, but the US and the UK were over-represented and France and Spain, under-represented. This may, however, have taken account of the geographical distribution of research a decade earlier rather than that of the late 1990s.

Increasingly, however, production and trade in the UK is subject to Directives and Regulations published by the European Commission. These cover an enormous range of products and services, and are either required to be enacted into law in individual Member States or are directly legally binding throughout the European Union. In the past, European law making has not been very transparent, but this is changing, and it is the Commission's intention to abide by revised standards of good practice regarding scientific expertise, including use of the results of research supported by the EU's Framework Programmes (European Commission, 2002). This will afford another route by which biomedical research can be seen to influence public policy.

6. NATIONAL AND REGIONAL NEWSPAPER ARTICLES

These documents can also be very helpful in bringing research news to a wider readership, and they involve quite different groups – politicians, senior officials, healthcare administrators and providers, other researchers and the general public. They suffer from the disadvantage in comparison with the documents described above in being 'unofficial', and subject to all sorts of biases in both selection and presentation. However they have the great merit of being rather easy to monitor and of incidentally allowing a good perspective on how research is viewed by the wider, non-technical, world. They can also be an important source of information for researchers. This was shown by Phillips et al. (1993). They found that papers in the *New England Journal of Medicine* that were published during a strike at *The New York Times* in the latter months of 1979, and were therefore not subject to journalists' reporting, were subsequently significantly less cited in scientific journals.

About two thirds of the articles in newspapers that report research advances cover what might be classified as biomedical research. A sample survey of two months' coverage of the UK national press in 2001 established the main parameters of biomedical research coverage in that country (Lewison, 2002). There were about 200 articles per month, and they were carried by all newspapers, not just the broadsheets (which are normally considered as more serious) — indeed, the *Daily Mail*, a tabloid, had the largest number of articles, as it makes a particular feature of health. One surprise of the study was the large number of journalists involved in writing articles: on some papers there would be a science correspondent, a medical correspondent, a social affairs correspondent, a legal correspondent, and even a defence correspondent all writing from time to time about research.

The emphasis of the articles was normally optimistic, with some enthusiasm for the results of the research, perhaps rather uncritical, although often opinions were sought on the context and significance of the research from representatives of medical research charities. (Not surprisingly, they tended to be somewhat sceptical and to emphasise that much more research was needed)! There was no evidence from this small survey that experiments on animals were regarded with anything other than respect, which may bode well for the climate in which researchers can work. However, questions were sometimes raised about ethical issues, particularly the use of foetal material and samples obtained from cadavers, and the provision of informed consent.

The methodology for the collection of newspaper citations is rather different from that used for clinical guidelines and other regulatory documents described above. In the first place, most newspaper articles refer to only one research article, although sometimes others are cited in order to provide context. It is important to record the normal bibliographic data on the citing article, such as newspaper name, date, page number, length of article, journalist(s) name — not always given — title and a brief synopsis. A subject classification is also needed; this and the newspaper name will appear as codes, probably three or four letters, to be taken from an extensive thesaurus.

It is sometimes a bit difficult to identify the cited research paper. It may be a journal article, often from one of the weeklies (*BMJ, Lancet, Nature, New Scientist, Science*) or it may be a paper presented to a conference. However, there are usually some clues such as the name of one of the authors, the cities and countries in which the research was done, or the journal name, so that a simple search should uncover the source item. Sometimes it is evident that two newspapers are writing about the same piece of research, and one of them gives additional details that can help to locate it.

The papers cited by the newspapers should be looked up and their bibliographic data recorded, including the formal reference (journal, year, volume, issue, pagination), names of all authors, title, addresses, and funding information. The latter can be codified with the use of the ROD thesaurus of biomedical research funding bodies, which currently numbers about 10,000 such organizations in most countries of the world. It is desirable for authors' names and addresses to be in a standard format, and for this purpose the SCI one is the most satisfactory, with its standardised address contractions and country names.

7. INTERNATIONAL CONSIDERATIONS

Each of these five tasks, and others, can be carried out in a single country in respect of its own publications, but the value of the data that are compiled as a result will increase greatly if corresponding work is also performed in other countries and the results combined. For this purpose it will be convenient if each database is maintained in a single centre and made accessible *via* the Web to other parties who have contributed data, or who have supported the work financially. There will need to be agreement between the various contributing parties on the exact format for the data held in each database and on how each contributor will be able to input their additional data, and gain access to that of others. This does not necessarily have to be through a direct input process; indeed it may be easier in the first instance if contributions are sent as e-mail attachments so that they can be checked for compliance with the data protocols before being added to the main file.

Once each database of cited papers, with information on the citing sources for each, has attained a certain size and international scope, it could be used for many types of international comparative studies. Some of these would be purely academic and for these it would, in principle, be desirable to allow free access. Others would be carried out for clients on a repayment basis and some part of the fee should be paid as a royalty to help with the costs of database maintenance and management. Some sponsors of research, or research performers, might wish to have regular reports on the extent to which their own work was being cited in this way as a means to its evaluation, and this would potentially provide a source of revenue for each of the databases.

ACKNOWLEDGEMENTS

The work on the identification and analysis of the references on the NICE and SIGN guidelines was carried out by Isla Rippon and Kate Wilcox–Jay. It was financially supported by The Wellcome Trust (the principal sponsor of the Research Outputs Database through a contract to City University), the Chief Scientist Office, Scottish Executive, and the Medical Research Council. The work on the identification and analysis of the references on the CAC pesticide evaluations was performed by Potenza Atoigbe as part of her work for an MSc in Information Science at City University. The work on newspaper citations was made possible by the compilation, *What the papers say*, produced by Elizabeth Graham and her colleagues in the Wellcome Library.

REFERENCES

Anderson, J., Williams, N., Seemungal, D., et al. (1996). Human genetic technology: exploring the links between science and innovation. *Technology Analysis and Strategic Management*, 8, 135–156.

Annandale, E., Hunt, K. (Eds). (2000). *Gender inequalities in health*. Buckingham: Open University Press.

Atiogbe, P. (2001). *Bibliometric analysis of the international codex alimentarius food pesticide standards*. MSc dissertation, Department of Information Science, City University, London.

Braun, T., Glänzel, W. (2000). Chemistry research in Eastern Central Europe (1992–1997). Facts and figures on publication output and citation impact for 13 countries. *Scientometrics*, 49, 187–214.

Chalmers, I., Sinclair, J.C. (1985). Promoting perinatal health — is it time for a change of emphasis in research? *Early Human Development*, 10, 171–191.

Colborn, T., Dumanoski, D, Myers, J.P. (1996). *Our stolen future — are we threatening our fertility, intelligence and survival?* New York, NY: Dutton.

Dawson, G., Lucocq, B., Cottrell, R., Lewison, G. (1998). *Mapping the landscape: national biomedical research outputs 1988–95*. London: The Wellcome Trust, Policy Report no 9.

Dean, M. (1999). A quiet clinical revolution begins. *The Lancet*, 353, 651.

Dean, T. (2002). The Cochrane collaboration and its contribution towards the management of allergic diseases *Clinical and Experimental Allergy*, 32, 1269–1273.

Doll, R., Peto, R. (1981). The causes of cancer — quantitative estimates of avoidable risks of cancer in the United States today. *Journal of the National Cancer Institute*, 66, 1191–1308.

European Commission (2002). *Communication from the Commission on the collection and use of expertise by the Commission*. COM(2002) 713 final, 11 December.

Garfield, E. (1955). A new dimension in documentation through association of ideas. *Science*, 122, 108–111.

Garfield, E. (1963). Citation indexes in sociological and historical research. *American Documentation*, 14, 289–291.

Garfield, E. (1979). *Citation indexing: its theory and application in science, technology and the humanities*. New York, NY: Wiley.

Glänzel, W. (2000). Science in Scandinavia: a bibliometric approach. *Scientometrics*, 48, 121–150.

Grant, J. (1999). Evaluating the outcomes of biomedical research on healthcare. *Research Evaluation*, 8, 33–38.

Grant, J., Cottrell, R., Cluzeau, F., Fawcett, G. (2000). Evaluating 'payback' on biomedical research from papers cited in clinical guidelines — applied bibliometric study. *BMJ*, 320, 1107–1111.

Grupp, H., Schmoch, U. (1999). Patent statistics in the age of globalization — new legal procedures, new analytical methods, new economic interpretation. *Research Policy*, 28, 37–396.

Hanney, S., Frame, I., Grant, J., Green P., Buxton, M. (2003). *From bench to bedside — tracing the payback forwards from basic or early clinical research — a preliminary exercise and proposals for a future study*. Brunel University, Health Economics Research Group report no 31.

Harhoff, D., Narin, F., Scherer, F.M., Vopel, K. (1999). Citation frequency and the value of patented inventions. *Review of Economic Statistics*, 81, 511–515.

Ingwersen, P., Larsen, B., Wormell, I. (2000). *Applying diachronic citation analysis to research program evaluations*. In B.Cronin, H.B. Atkins (Eds), The Web of Knowledge — a Festschrift in Honor of Eugene Garfield (pp 373–387). Medford, NJ: Information Today, Inc.

Kostoff, R.N. (1998). The use and misuse of citation analysis in *Research Evaluation* — comments on theories of citation. *Scientometrics*, 43, 27–43.

Laupacis, A. (2002). The Cochrane Collaboration — how is it progresssing? *Statistics in Medicine*, 21, 2815–2822.

Lederberg, J. (2000). How the science citation index got started. In B.Cronin, H.B. Atkins (Eds), *The web of knowledge — a Festschrift in honor of Eugene Garfield* (pp 25–64). Medford, NJ: Information Today, Inc.

Lewison, G. (2002). From biomedical research to health improvement. *Scientometrics*, 54, 229–240.

Lewison, G. (2003). Austrian biomedical research — a bibliometric evaluation. *Plattform Forschungs- und Technologieevaluierung*, 18, 13–17 – see Figure 5.

Lewison, G., Paraje, G. (2003). *The classification of biomedical journals by research level*. Proceedings of the 9th International Conference on Scientometrics and Informetrics, Beijing, China; pp 142–151. Also *Scientometrics* (2004), in press.

Lewison, G., Wilcox–Jay, K. (2003). *Getting biomedical research into practice — the citations from UK clinical guidelines.* Proceedings of the 9th International Conference on Scientometrics and Informetrics, Beijing, China; pp 152–160.

Leydesdorff, L. (1998). Theories of citation. *Scientometrics*, 43, 5–25.

Mahmoud, A.A.F. (Ed). (2001). *Schistosomiasis*. London: Imperial College Press.

McGinnis, J.M., Foege, W.H. (1993). Actual causes of death in the United States. *Journal of the American Medical Association*, 270, 2207–2212.

Moed, H.F. (2000). Bibliometric indicators reflect publication and management strategies. *Scientometrics*, 47, 323–346.

Moxham, H., Anderson, J. (1992). Peer review: a view from the inside. *Science and Technology Policy*, Feb, 2–6.

Narin, F. (1994). Patent bibliometrics. *Scientometrics*, 30, 147–155.

Narin, F., Olivastro, D. (1998). Linkage between patents and papers: an interim EPO/US comparison. *Scientometrics*, 41, 51–59.

Noyons, E.C.M., Luwel, M., Moed, H.F. (1998). Assessment of Flemish research and development in the field of information technology — a bibliometric evaluation based on publication and patent data, combined with OECD research input statistics. *Research Policy*, 27, 285–300.

Peritz, B.C. (1992). On the objectives of citation analysis – problems of theory and method. *Journal of the American Society for Information Science*, 43, 448–451.

Petrie, J.C, Grimshaw, J.M., Bryson, A (1995). The Scottish Intercollegiate Guidelines Network Initiative — getting validated guidelines into local practice. *Health Bulletin*, 53, 345–348.

Chapter 21

THE FOUR LITERATURES OF SOCIAL SCIENCE

Diana Hicks

School of Public Policy, Georgia Institute of Technolog, GA, USA
E-mail: diana.hicks@pubpolicy.gatech.edu

Abstract: This chapter reviews bibliometric studies of the social sciences and humanities. SSCI bibliometrics will work reasonably well in economics and psychology, whose literatures share many characteristics with science, and less well in sociology, characterised by a typical social science literature. The premise of the chapter is that quantitative evaluation of research output faces severe methodological difficulties in fields whose literature differs in nature from scientific literature. Bibliometric evaluations are based on international journal literature indexed in the SSCI, but social scientists also publish books, write for national journals and for the non-scholarly press. These literatures form distinct, yet partially overlapping worlds, each serving a different purpose. For example, national journals communicate with a local scholarly community, and the non-scholarly press represents research in interaction with contexts of application. Each literature is more trans-disciplinary than its scientific counterpart, which itself poses methodological challenges. The nature and role of each of the literatures will be explored here, and the chapter will argue that by ignoring the three other literatures of social science bibliometric evaluation produces a distorted picture of social science fields.

1. INTRODUCTION

Bibliometrics has proved a powerful tool for the evaluation of scientific research. The application of bibliometric method to research in disciplinary areas in which consensus is reached has become almost routine. Bibliometric work is facilitated in such areas because their literature exhibits certain characteristics: research is published predominantly in English language journals and references predominantly recent papers in a set of core journals

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 473-496.

recognised for their high quality and impact. Thus, a focused body of citations is generated which is fairly current and is accessible if a bounded set of journals is indexed. The *Science Citation Index* of course takes advantage of these characteristics to provide the indispensable basis for citation analysis of scientific output. If research outcomes are to be evaluated, patent citations to scientific literature are available (Narin, 1997), and these are almost as well indexed and well behaved as the journal literature. They are also becoming more useful as more and more public sector researchers patent (Hicks et al., 2001).

When challenged to evaluate scholarly work in the social sciences and humanities, we are rudely forced to work outside this comfort zone in a frankly messy set of literature. In the humanities book publishing predominates, and even today books and their references are not indexed in a database. In the social sciences indexed English language journal publication coexists with non-indexed book publishing, national literature, and non-scholarly literature. In the humanities referencing is archival (de Solla Price, 1970) and citations accumulate at a geological pace from the perspective of policy makers. In the social sciences referencing mixes archival and current patterns and the referencing pattern is quite scattered, lacking focus. A core literature is less clearly delineated.

This chapter will interpret the situation within the Mode 2 framework. Mode 2 is the simplification of the argument first put forth by Gibbons et al., namely that:

> "The old paradigm of scientific discovery ('Mode 1') characterised by the hegemony of disciplinary science, with its strong sense of an internal hierarchy between the disciplines and driven by the autonomy of scientists and their host institutions, the universities, was being superseded — although not replaced — by a new paradigm of knowledge production ('Mode 2') which was socially distributed, application–oriented, trans-disciplinary, and subject to multiple accountabilities"(Nowotny, Scott and Gibbons, 2003, p. 1).

Nowotny, Scott and Gibbons (2001) note the pervasiveness of processes of audit, assessment and evaluation in Mode 2. Bibliometrics has proved remarkably adept in implementing this agenda in the sciences. Thus bibliometrics is asked to extend itself into social science and humanities. Ironically, this tool of the Mode 2 'audit culture' works best on traditional Mode 1 science areas.

In confronting the social sciences in particular, I will argue that bibliometrics confronts evaluating aspects of Mode 2 research. The chapter examines the four literatures of social science: journal article, books, national, and non-scholarly literature. The discussion explores their relationship to scientific and humanities scholarship and to trans-disciplinarity and contexts of application. The chapter will examine the methodological problems of the four literatures and will assess the success of efforts to resolve the problems and the consequences of ignoring them.

Note that 'social science' or 'humanities' will not be analysed here because generalisations at that level are of limited use. The bibliometric literature takes a more nuanced approach, examining issues at the field level, which has proved valuable. In almost every study the psychology and economics literatures are found to be most science–like, in contrast with the sociology literature. Also fields change over time. Zwaan and Nederhof (1990) point out that some parts of linguistics have converged towards cognitive science and publication patterns have come to resemble social sciences more than history. Thus core journals can be identified and the average reference has become more recent. Bibliometrics becomes quite tractable, even in this area traditionally viewed as a humanities field. We should beware of very old studies, as their results may not reflect the current situation.

2. JOURNAL ARTICLES

The first literature of social science comprises internationally oriented, largely English language, peer reviewed journal articles. The SSCI indexes these, enabling evaluations applying classic bibliometric technique whose authors acknowledge to varying degrees their exclusion of the three other literatures.

Glänzel (1996) worked from the full SSCI database to produce tables listing countries' publication and citation counts and shares, and citation per paper indices between 1990 and 1992. Glänzel recognised the substantive methodological problems arising from the nature of the social science literature, and proposed that his SSCI based indicators be interpreted cautiously. In his methodological work Glänzel has devoted considerable attention to the time distribution of citations. In this evaluation he was forced to acknowledge that although a decade long citation window would be needed to capture the slow accumulation of citations in social science, from the evaluation perspective, such methodological rigor would produce an obsolete result. Glänzel compromised with a shorter window and as a result, compared to SCI data, "mean citation rates are ... small, and the share of uncited literature is considerable" (Glänzel, 1996, p. 293).

Ingwersen, in a series of papers, examines at the national level Scandinavian publication and impact in social science and medical areas. Ingwersen begins with on-line publication counts and later moves to the

ISI's National Science Indicators product (NSI) containing national level summary publication counts. All of the papers compare Nordic countries with the world and with each other in publication output and citation impact by field within social science or health areas. The countries tend to produce high impact work in the health sciences and each has individual strengths in social science areas. In some cases trends and strengths could be connected with policy, for example the connection between strong social welfare states and strength in health sciences. As for methodological awareness, Ingwersen 2000, an NSI based analysis of traditional social science fields, finds that Scandinavian output is increasing and in many cases a country's share of ISI literature is comparable to their share of scientific literature. Whilst admitting to the continuing Anglo–American bias of the database, Ingwersen concludes that increased publication output by small countries in the SSCI makes it increasingly relevant for analysis of non-US countries in five to seven of the nine fields examined. (Ingwersen, 1997, 2000, 2002; Ingwersen and Wormell, 1999)

Katz (1999) worked from the NSI to compare national levels of social science journal publication. The UK was the focus, and Katz found that the UK share of papers increased between 1981 and 1998. Larger and faster growing fields were identified for the UK and its constituent regions. In examining citations Katz argued that a linear normalisation, i.e. citations per paper, is inadequate because citation counts increase non-linearly with size of the publication pool. He introduced a corrected indicator more favourable for small countries. On methodological issues Katz incorporated much of Hicks 1999 to conclude: "bibliometric indicators may provide a reasonable measure of the size and impact of international and scholarly social science research in some fields like psychology and economics" (p. 4). The report focused on psychology and economics.

Godin (2002) works from the full database. He counted Canadian papers by province, by sector, and by field, and counts collaborations at the sector level. He identified health and psychology as areas of Canadian specialisation. Aligned with Ingwersen, Godin noted that Canada's share of papers in the social sciences stands at 5.8%, larger than its share of papers in the sciences and engineering — which is slightly over 4%. This was seen as evidence that the SSCI was useful for social science evaluation.

The most detailed and methodologically careful evaluations of social science and humanities research have been undertaken by the Leiden group, Nederhof in particular. The group's work has been guided by conversations with topic experts, methodological issues were always acknowledged, and the analysis has been deeper than is typical elsewhere.

In the late 1990s Nederhof and Van Wijk mapped social and behavioral science topics and disciplines using the SSCI. They generated maps by clustering a matrix whose rows listed topics (title words) and whose columns listed disciplines (a consolidation of ISI's journal classification scheme). Two maps were analysed, a dynamic and a static map. The dynamic map was built using words whose frequency changed greatly. The static map was built from the 100 most frequently occurring non-trivial words. In one paper the authors examined Dutch areas of strength and weakness. They found that Dutch performance had some strong areas, but was slightly disappointing overall (Nederhof and Van Wijk, 1997). In another publication the authors dug deeper into the maps to profile Dutch institutes. This necessitated adding back into the analysis topics missed in the quite selective mapping process. The results were quite complex and suffered from thin citation – in some cases a well cited output had two citations (Nederhof and Van Wijk, 1999).

All these SSCI–based evaluations handled the SSCI data well. They produced useful insights into national patterns of publication in SSCI–indexed journals. The authors also acknowledged the methodological issues inherent in SSCI–based bibliometrics. Nevertheless, a problem lurks behind these evaluations: social scientists publish in more than just SSCI–indexed journal articles. Bourke, Butler, and Biglia examined two bibliographies of Australian university research output. They found that natural scientists published about 85% of the time in journal articles or published conference papers; whilst for social scientists and the humanities the figure was about 61%. Books, edited books, book chapters, monographs and reports, creative works and 'other' accounted for the rest (Bourke et al., 1996). Pestaña, Gómez, Fernández, Zulueta, and Méndez examined Annual Reports to construct a bibliography of the research output of the Spanish Scientific Research Council (CSIC). The CSIC's seven natural science divisions published 81% of their output in journals and the one humanities/social science division 54% (Pestaña et al., 1995). Winterhager has examined German sociology publishing in the German SOLIS database and found that 42% is published in journals (Winterhager, 1994). Thus journal–based bibliometric indicators will be based on a smaller fraction of research output in the social sciences than in the natural sciences.

Luwel et al. took this point very seriously in his project analysing the research activities of four major Flemish universities in law and linguistics. The study included no citation information nor did it draw evaluative comparisons amongst the universities. Rather the study represented an extended discussion with representatives from the law and linguistics faculties in the four universities with input from publication counts. Based on survey data, the authors analysed how scholars spent their time, turnover rate amongst scholars, complex self–reported sub-disciplinary structures, external funding, prizes, and publications classified into 30 categories. Surveys also gathered information on peer recognition of scholars and local and international impact of journals. The authors' extensive cleanup and classification of research output combined with their rating of journals for international impact and quality provided a means of devising output indicators independently of the SSCI that overcome some of the methodological concerns haunting pure SSCI work (Luwel et al., 1999; Nederhof et al., 2001; Moed et al., 2002).

Lewison in 2001 also addressed these concerns in his evaluation of UK output in a humanities field — the history of medicine — that focused on assessing book output. Lists of books in the history of medicine were compiled from book reviews and from references in papers listed in the SSCI. Author addresses were gathered from SSCI papers and one-quarter of the books could be assigned to countries in this way. Citations in the SSCI and book reviews (indexed in the SSCI and so easily accessible) were counted. The UK was found to be increasingly strong in the field, supporting the results of an international series of interviews. Methodologically the counts of reviews and citations did not correlate, in fact there was little overlap between books that were reviewed and cited. When asked, historians responded that reading a book is the best way of evaluating it, followed by reading a review and then by the number of citations. The number of reviews ranked considerably lower on the list.

Non-journal publishing is significant in the social sciences. Some have wrestled with this problem; others have acknowledged it. In addition to non-journal publishing a second factor compromises SSCI–based evaluations — the robust trans-disciplinarity of much social science. The bibliometric evidence for this trans-disciplinarity is found in widely scattered citation patterns. Beginning at the broadest level, Leydesdorff reports that 79% of references from papers indexed in the SCI are references to other papers indexed in the SCI. In contrast, 45% of references from papers indexed in the SSCI are within the database (Leydesdorff, 2003).

Small and Crane (1979) conducted a co-citation clustering of high–energy physics, psychology, economics, and sociology 1972–1974 using the full SCI and SSCI. Examining the characteristics of the resulting clusters, they found strong evidence of trans-disciplinarity in sociology compared with the other areas. For example, 97% (all but one) of the sociology clusters was considered interdisciplinary in that less than 2/3 of the citing papers were in sociology journals. In contrast, in psychology and economics a smaller proportion of the clusters were interdisciplinary using the same criterion (71% and 64% respectively). Examining co-citation links between clusters in the disciplines revealed that economics clusters were substantially more strongly linked to each other than were the sociology clusters. Examining links between clusters and other disciplines revealed that sociology clusters have more connections with other fields than do economics clusters. Small and Crane's work revealed that in comparison with economics, sociology's citing patterns were less focused on literature in the same field. Sociology clusters were less strongly linked to each other and more strongly linked to clusters in other fields. Thus sociology was more trans-disciplinary than economics.

Similar evidence of trans-disciplinarity emerged from a study by Glänzel et al. (1999). These authors also analysed references in the SSCI, using them to attempt to classify papers based on the subject classification of journals they referenced. The technique aimed to classify papers in journals selectively covered by the SSCI, which are not assigned to fields. The authors counted references to journals which had been classified into business, economics, law, political science, psychology, sociology, or information and library science. The field referenced most often was used as the new classification of the paper if its share of references exceeded 50%. If there were no references to these fields, or no field gathered 50% of the references, the paper could not be classified. In all, 28% of the papers could be assigned to a social science field. That 70% of papers could not be classified speaks to their trans-disciplinary nature. Interestingly, the method was also applied to two disciplinary journals. 25% of the papers in the American Sociological Review (ASR) could not be classified as sociology, whilst 6% of papers in Developmental Psychology could not be assigned to psychology. Sociology again appears more trans-disciplinary than a comparison field, in this case psychology.

Broad, unfocused citing fragments the literature so that in the worst cases no core of literature in a field can be identified (Nederhof et al., 1989). A database such as the SSCI must have an internationally recognised core literature to work with to achieve comprehensive international coverage. Low SSCI coverage of a journal literature may signal no core literature. We might expect fragmentation to vary by field, and less trans-disciplinary fields to be the least fragmented, and so it is not surprising to find that SSCI coverage varies by field, with economics and psychology literature the best covered.

Two studies provide detailed field breakdowns of their coverage figures. Table 21.1 reports Nederhof et al.'s (1989) finding that coverage of Dutch output ranged from 62% of journal articles in experimental psychology to 2% in public administration. Table 21.2 reports Butler's findings (personal communication of unpublished data, 1998) that coverage of Australian anthropology, archaeology, philosophy, law, and economics was more than 40%. In contrast, only 25% of history was covered. In Butler's data there was an inverse correlation (minus 0.83) between share of journal articles indexed in the SSCI and share of total publications accounted for by books or chapters in edited books. That is, the more books in a field, the smaller the share of its Australian journal literature covered by the SSCI.

Table 21.1. SSCI coverage by field – Nederhof (1989) Dutch Social Science

Field	*% of articles in SSCI*	*% publications in books*
Experimental Psychology	62	30
General Linguistics	21	40
Anthropology	15	38
Dutch Language	10	25
Social History	10	40
Public Administration	2	36

Butler's result extends the trans-disciplinary argument by linking a lack of core literature and the presence of many books. If trans-disciplinarity varies by field then fields with a higher share of books according to Pierce (1987) should have less core journal literature according to Nederhof et al. (1989). In Butler's data economics, and anthropology and archaeology exhibited the highest share of articles covered and a low share of books while history exhibited the opposite pattern.

Table 21.2. SSCI coverage by field – Butler (1998), Australian social science

Field	*Number of articles*	*% articles in SSCI*	*% publications in books*
Anthropology & Archaeology	281	44	6
Economics	1,074	43	4
Philosophy & Law	418	43	8
Geography	390	39	5
Sociology	649	32	9
Political Science	690	27	8
Asian History	220	27	10
History	532	25	12
Total	4254	35	7

3. BOOKS

The second literature of social science is books. The association between books and trans-disciplinarity is supported by citation evidence. In 1971 Broadus surveyed the literature of citation studies in the social sciences and found 11 studies, 6 of which used books (technically monographs) as sources of citations. He found evidence that books referenced more widely than journal articles. That is, in comparison to a journal article, a higher percentage of references from a book will be to work outside its specialty

(Broadus, 1971, p. 238). Looking at citations gathered by books, Clemens' et al. studied sociology and reported that books received the majority of citations from outside the discipline of sociology. In the least cited quartile books received 54.5% of their citations from outside sociology compared with 16% of citations to journal articles. In the most cited quartile books received 79% of their citations from outside sociology and articles 55%.

The trans-disciplinarity of books suggests that the book and journal literatures differ, a point pursued further below. However, books are a small percentage of social science output, and so one might choose to ignore them. The reason one cannot is that books have a high impact in social science. Broadus' review found that references to monographs ranged from 31% to 56% of references from book and journal literature in a variety of fields. He compared this with a 1939 study showing chemists gave 5% of their references to monographs and physicists 8% (Broadus, 1971, p. 241). Small and Crane (1979) analysed references from journal articles indexed in the SCI and SSCI and found that the share of the cited items that were books was:

– 0.9% in high energy physics;
– 15% in psychology;
– 25% in economics;
– 39% in sociology.

Thus books are ignored in studies of science, but in social science, although a relatively small percentage of output, they account for a substantial proportion of citations in the SSCI — as much as 40%. Indicators built from SSCI indexed material — journal articles and citations to them — will miss the 40% of citations received by books. Books can be very highly cited:

– Hicks and Potter (1991) examined a bibliography of sociology of scientific knowledge and found that on average journal articles received 1.2 citations and books 5.7 citations.
– Clemens et al. (1995) compares the citation rate of elite publications: papers published in the two leading American sociology journals — American Sociological Review and American Journal of Sociology in 1987 and 1988 — and 80 books nominated for the American Sociological Association's Distinguished Scholarly Publication award. They find that "books are clearly cited more frequently than journal articles by a ration of 3:1" (p. 459). Citations to the 20 most cited articles ranged from 16 to 55 while citations to the 20 most cited books ranged from 34 to 512.

- Bourke et al. (1996) examined research output 1989 to 1993 for social sciences at the Australian National University and found that on average journal articles received 0.9 citations and books 5.2 citations.
- Thomas (1998) collected a bibliography of 300 items published by leading authors in organisational behaviour between 1956 and 1975. The 33 most cited items were books.
- Webster's (1998) lists of most cited Polish sociology documents are mostly books — 11 out of 15 cited in the SSCI and 18 out of 19 cited in the Polish Sociology Citation Index.

This evidence establishes that books are high impact, and thus under the rules of bibliometrics should not be ignored. The danger of ignoring books is further illustrated by exploring the differences between the worlds of book and journal publishing. Books are not just large journal articles. Evidence is found in the lack of correlation between cites to books and journal articles. Four studies illustrate these points:

- Nederhof et. al. (1989) lists the citations per book and journal article for 19 departments; the correlation between the two was 0.32.
- Hicks and Potter (1991) collected a bibliography of 17 authors' output in the field of sociology of scientific knowledge. The correlation coefficient of the citation per book and journal article figures was 0.35.
- Bourke et al. (1996) compared the rankings of departments using total and journal only citation counts They concluded: "In the social sciences and humanities, the use of journal citation rates as a surrogate for total publication citation rates is more likely to be misleading than in the sciences. It still does, however, provide useful information when used in conjunction with informed peer review" (Bourke et al., 1996, 54).
- More recently, Cronin et al. (1997) constructed a database comprising 30,000 references from 90 books randomly chosen from those reviewed in top sociology journals and published between 1985 and 1993. Cronin et al. compared lists of the 26 authors most cited in the monographs and in the top 24 sociology journals. They found that nine authors featured on both lists. The five authors ranked 22 to 26 on the book list did not appear among the top 532 authors most cited in the journals.

The low correlations in citation counts combined with the differing highly cited author sets suggests that the journal and book literature form different worlds. That these worlds may overlap but retain a distinct identity is supported by Line's work. Line constructed a set of 59,000 references, 11,041 from monographs and 47,925 from journals. Line found that, compared to journals, monographs referenced proportionally fewer journal articles and more monographs and other types of literature. This suggests that the journal and book literatures are somewhat self–contained, although obviously interdependent and overlapping.

Table 21.3. References made by journals and monographs to other forms of material (source: Line, 1979, p. 274)

Forms of material cited	*Source material* *Journal articles*	*Monographs*
Journal articles	47%	25%
Monographs	39%	51%
Other newspapers, unpublished etc.)	14%	24%
Total	100%	100%

The different types of scholarship they represent may explain why two worlds of literature coexist. Journal articles may reflect a more scientific, and books a more humanities type of approach to scholarship. Clemens et al.'s study of sociology helps us understand this. Clemens et al. compared book and journal publishing within the context of a long standing debate in sociology. Is the field professional, technical, cumulative, and convergent as one would gather from its journal literature or is it a diversified, intellectually open endeavour as found in the books? Examining the two types of publishing sheds light on the themes of scientific integrity versus intellectual vitality that underpin the debate.

Clemens et al.'s evidence supported the notion that book and journal publishing form different worlds. Entry into article publishing, they argued, is competitive and so more egalitarian than entry into book publishing, which relies more heavily on patronage, recommendations and reputation. Thus they found that book authors were more likely to be trained and located at elite private universities than were journal article authors. Article authors were more junior than book authors. Articles were more likely to be based upon quantitative evidence and books on qualitative evidence (although books based on quantitative evidence were the most cited of all). They concluded:

> "... books and articles play different roles. Books are high–stakes endeavours that, when successful, are effective in enrolling allies from neighbouring fields. Articles, in contrast, discipline the troops, generating a common currency of evaluation, be it in comprehensive exams or tenure decisions. To the extent that we care about scholarly reputation, both our discipline's and our own, neither genre should be ignored"(Clemens et al., 1995, p. 484).

Clemens et al.'s analysis painted a picture of a heterogeneous field of scholarship with distinct journal and book traditions. Journals represent a more scientific type of research and books a more humanities type of scholarship. Both are trans-disciplinary, books more so. Because books are very highly cited and often produced by different people than journal articles, SSCI–based analyses will differ from more inclusive studies. Bibliometricians ignoring books risk distorting our picture of social science.

4. NATIONAL LITERATURES

The third literature of social science is national. American and European geologists are interested in Iceland's volcanoes, and geneticists learn much from Iceland's genealogical records (Thorsteinsdottir, 1998), but Dutch journals in public administration remain unknown to foreign experts (Nederhof, 1989, p. 338). In contrast to science, social sciences are more embedded in their social context because society is their concern. Social science research agendas are influenced by national trends and by policy concerns of the national government. Theoretical concepts are subtle, and without the unifying language of mathematics are expressed in national languages, and can often be fully appreciated only in the original language. Countering this, Nederhof argues that:

> "Genuine scholarly research, regardless of the sub-discipline and the object of research, leads to results the relevance and implications of which go beyond a purely national viewpoint or interest. This may be less so for contributions of a more applied or practical nature. Therefore [at least some] outcomes of genuine scholarly research, even those primarily related to national aspects, deserve to be communicated — in an appropriate form — to scholars in other countries as well" (Nederhof, 1989, p. 513).

This section examines the existence and nature of national literatures. Here national and international literatures are juxtaposed. National journals are those which are not often indexed in the SSCI; which primarily, though not exclusively, publish articles in the native language (not English) of their country of publication, and whose authors and readers largely work in that country. International journals include most journals indexed in the SSCI (although parochial US and UK journals are often SSCI indexed); and are largely English language journals whose authors and readers work in many countries.

Bibliometric evidence suggests that both producers and consumers of social science are nationally oriented. Research shows that compared to natural scientists, social scientists both write for and read fewer foreign language or even foreign journals. Kyvik studying the writing habits of

Norwegian scientists and social scientists in the early 1980's, found that compared to the scientists fewer social scientists published in a foreign language and more published in Norwegian (Kyvik, 1988, p. 165). Taking authors' citation patterns as an indication of their reading habits, Yitzhaki (1998) found that authors over–cite material in their own language. American and British authors cited English language material 99% of the time, although English language sociology probably accounted for 70% of the world literature. German and French authors cited material in their own language more than 60% of the time although such material accounted for less than 10% of literature in the field. However, Nederhof et al. (1989) emphasised that visibility depends less on writing in the English language than it does on publishing in an international journal. That is, the impact of English language papers in Dutch journals is not higher than the impact of other papers in Dutch journals. In a sense then, each national literature is a world unto itself.

In addition, a national literature constitutes a world overlapping to a limited extent with the SSCI as was well illustrated by Webster/ Winclawska's analysis of a Polish sociological citation index (PSCI) (Webster, 1998; Winclawska, 1996). In the first analysis Winclawska began with a list of Polish sociologists and counted their citations in the international SSCI and the Polish index between 1980 and 1988. She found that of the top 10 most cited journals in the Polish index, only the three foreign ones are indexed in the SSCI.

In the second analysis the author, now Webster, counted citations to Polish sociologists between 1981 and 1995. She found:

– Lists of the top 20 most cited Polish sociologists in each index had 12 names in common. The most cited sociologist on the Polish list (with 253 citations) was ranked 41st in the SSCI (with 19 citations). The most cited sociologist on the SSCI list (with 254 citations) was ranked 20th on the PSCI list (with 41 citations).
– Lists of the top 20 most cited documents by Polish sociologists in each index contained none in common. All but one of the SSCI cited documents were in English; all the PSCI cited documents were in Polish.

The Webster/Winclawska's analyses illustrated the bibliometric consequences of the limited overlap between national and SSCI literatures. Bibliometric indicators based on foreign literature painted one picture of Polish sociology, and the Polish sociology index another.

Maintaining a database is far more demanding than compiling a list, and so database coverage can be compared against more comprehensive worldwide journal lists. Schoepflin (1990) compared the UNESCO 1986 World List of Social Science Periodicals with the list of journals indexed in the SSCI. Table 21.4 below is taken from Schoepflin's article. It compares the number of journals produced in the US, UK, Germany, and France that appear on the UNESCO list and in the SSCI. At that time UNESCO's list at 3,515 journals was 2½ times as long as SSCI's at 1,417. Interestingly, SSCI indexed more American journals than UNESCO, confirming the comprehensiveness of US coverage in the SSCI. The UK is also over–represented in the SSCI at 18%. German and French literature is not as well covered in the SSCI, nor is the rest of the world. Schoepflin's work confirms that except for the US and probably the UK, the SSCI and national literatures represent partially overlapping yet different worlds.

Table 21.4. Comparison of SSCI and UNESCO journal lists

	Number of Journals			*Percentage share*		
Country	SSCI		UNESCO	SSCI		UNESCO
USA	852	>	611	60	>	17
UK	256	<	334	18	>	10
Germany	48	<	184	3	<	5
France	25	<	269	2	<	8
Rest of world	236	<	2,117	17	<	60
Total	1,417	<	3,515	100	=	100

The proportion of a nation's output accounted for in indicators will depend not only on the number of a nation's journals indexed in the SSCI; it will also depend on how often researchers publish in English language international journals. Determining the share of national output indexed in the SSCI is laborious, nevertheless a variety of studies have examined this. Table 21.5 summarises the relevant parts of these studies, presenting the percentage of social science journal output indexed in the SSCI for a variety of countries.

There is quite a range in the figures. UK economics seems well covered with 73% of its articles indexed (Nederhof and Van Raan, 1993). This accords with Shoepflin's analysis, which showed UK journals are relatively well covered. About one-third of Australian and Dutch social science journal output is covered (Butler, 1998; Tijssen et al., 1996; Royle and Over, 1994), and a small percentage of Spanish output (Pestaña et al., 1992; Villagrá Rubio, 1992). Apparently the Spanish publish much more in Spanish than the Dutch do in Dutch.

Except for the US and UK, national social science literatures are largely excluded from the SSCI. SSCI indicators will represent internationally oriented research. Webster summarises this point well, concluding that the SSCI indicates the presence and the impact of Polish sociology on the international arena, focusing on areas of research done in Poland which are of interest to the international community and the 'best' Polish sociologists and Polish sociological works; but the SSCI "does not allow for an in-depth analysis of the local dimensions of the discipline" (Webster, 1998, p. 31).

Table 21.5. SSCI article coverage

Study	*Country (number of country's journals indexed in SSCI)*	*Number of journal articles*	*% of journal articles in SSCI*	*% of all publications in SSCI*
Nederhof 93	UK (278) – economics only	193	73	27
Burnhill	UK (278)	468	46	22
Butler	Australia (20)	4,254	35	
Tijssen	Netherlands (83 - 3 Dutch)	all Dutch[1]	30	
Royle & Over	Australia (20)	1,901	27[2]	
Pestaña	Spain (3)	1,242	4	2
Villagra Rubio	Spain (3)	3,757	1[3]	1
Winterhager	Germany (52)	49,446		25

[1] Elsevier English language journals are attributed to the Netherlands.
[2] Comparable figure for science: 74% of 6304 articles indexed in SCI.
[3] Strictly speaking this is percentage in 'international journals', i.e. those indexed in any of 11 international databases including Social Scisearch.

However, the prospects for social science indicators may be improving as social scientists become more internationally oriented. There is some bibliometric evidence on this point from the studies reviewed here:

– Pestaña et al. (1995) mention that the Spanish CSIC research output is growing more international, though they do not say if this trend is strong in the social sciences sections;
– Van der Meulen and Leydesdorff found that the proportion of Dutch philosopher's output published in foreign, scholarly journals increased from 3% to 17% between 1979–80 and 1984–85 (Van der Meulen and Leydesdorff, 1991, p. 309).

There are clearly forces working towards the homogenisation of social sciences — economic globalisation; the internet; European research funding that requires international collaboration; the transitions of East and Central European nations that freed communication and travel, and national level evaluations that emphasise publishing in high impact journals (such as the UK Research Assessment Exercise).

In fact, in Nederhof and Van Wijk's (1997) word–based topic clustering in the late 1980's (described earlier) the authors found that in the international literature indexed in the SSCI:

With the exception of a minority of topics related to political science, to social issues, and to a lesser extent physical health and geographical location, the large majority of the topics seem to reflect a transnational substantive interest. In addition, the [US and European countries] studied here share many social and political issues. Of course, this may not be true for other countries, and in particular non-Western countries. The present data suggest that the research front on many topics in the social and behavioural sciences is international in the late 1980s. Of course, this does not preclude that publications on national issues or national aspects of issues appear in journals or books that address primarily a national audience (p. 271).

Perhaps the most intriguing evidence on increasing internationalisation of social science, and hence of the SSCI, is provided by comparing the Winclawska and Webster studies. Her first study covered pre-transition Polish sociology, 1980 to 1988, her second covered pre and post transition sociology. Pre-transition, the SSCI missed 90% of Polish sociologists; post transition, it missed only 30% — a figure much closer to the Polish Sociology Citation Index (PSCI).

The quantitative evidence suggests that the overlap between the worlds of national literatures and the SSCI has increased. At the same time the continued existence and differentiation of national literatures is not in question. Note the heavy caveats on Nederhof and Van Wijk's statement above; in addition Webster's work added nuance to the argument. Webster's work suggested that the ascendancy of an international social science may place small–country social scientists in the position of applying other's frameworks to their societies, recognised internationally mostly when their societies present picturesque episodes that become fashionable topics in big countries. National communities may develop method and theory, but big–country social scientists remain impervious.

This conclusion was suggested by comparing the topics of the works most highly cited in the PSCI and SSCI. Polish sociologists highly cited (in articles published in the four Polish journals indexed in the PSCI) handbooks in general sociology by Polish authors, works on the social structure of Polish society, and works on interesting theoretical or methodological issues. Works highly cited in the SSCI included: 6 dealing with theoretical issues, each at least 20 years old; and the rest dealing with social unrest in Poland in the early 1980s and the fall of Communism in Eastern Europe. Webster concluded that: "the international sociological community does not notice Polish attempts to tackle universal issues in sociology; it is primarily interested in 'fashionable' topics and fads associated with the 'velvet revolution' and systemic transformation." (Webster, 1998, pp. 23–24).

Small country social scientists can be internationally recognised, but perhaps have fewer possible strategies for doing so than US or UK social scientists. Many may choose to pursue topics which will not interest those in other countries. National literatures will provide a more complete picture of many social science fields in small countries because they will include theoretical and methodological development. Increasing internationalisation may thus work to change the nature of social science in small countries. Ingwersen argued that analysis is possible when the number of a country's papers in a social science field that are indexed in the SSCI becomes reasonable, i.e. as the country's share of world output in the social science field approaches its share in scientific fields. However, as with books, what is missed is not the same as what is counted. One world is delineated; another exists.

5. NON-SCHOLARLY LITERATURE

The fourth literature of social science comprises non-scholarly works. Non-scholarly journals are those "usually directed at non-specialists such as high school teachers or, in short, the general public . . ." They are devoted to enlightenment or knowledge transfer to the non-scholarly public (Nederhof and Zwaan, 1991, p. 335). In the US the economist Paul Krugman exerts influence through his New York Times column. Burnhill and Tubby-Hille found that in the UK "projects in education [were] reaching practitioners through such periodicals as the Times Education Supplement, with researchers in sociology, social administration, and socio-legal studies publishing in such periodicals as New Society and Nursing Times". (Burnhill and Tubby-Hille, 1994, p. 142) Where national literatures can develop knowledge in the context of application, publishing in non-scholarly journals moves knowledge into application. The literature therefore performs a function similar to patenting for scientists. But patent systems are indexed, can contain citation structures amenable to bibliometric analysis, and have gained respect as a valued output worthy of evaluation (Narin, 1994). In contrast, non-scholarly literature, being also national literature, is less well indexed, does not earn citations and has not yet earned respect as a valued output of scholarly work interacting with application.

Burnhill and Tubby-Hille (1994) have investigated this issue in some depth. Their publications database was constructed from end–of–award reports of grant holders to the granting agency, supplemented by a survey. They checked whether listed journals were peer–reviewed using two directories of periodicals which identify peer–reviewed serials – EBSCO and Ulrich's. Burnhill and Tubby-Hille then examined SSCI coverage of 'peer–reviewed' journals. The SSCI indexed 82% of articles in journals regarded as peer–reviewed by the directories or at least two authors. However, the

SSCI coverage dropped to 67% if articles in self–reported 'scholarly' journals were included.

Burnhill and Tubby-Hille did not report SSCI coverage by field. However, they did report scholarliness of articles by field (Table 21.6). In this table, 'peer–reviewed' means articles in journals judged to be peer reviewed by the directories or by two or more authors. 'Authors consider scholarly', means an author reported the article to have been peer reviewed on the survey. 'Other' is remaining journal articles. Psychologists, statisticians and geographers do not publish much in non-scholarly literature. Other fields do. Economics here diverges from its more general pattern of scientific type publishing with a healthy percentage of articles in non-scholarly venues. Linguistics, education and sociology lead in share of non-scholarly publications.

Nederhof et al. (1991) have also looked quite closely at this issue. They surveyed Dutch and foreign scholars asking them about the scholarliness of a number of journals in which Dutch social scientists published. They found that journals considered scholarly in university annual reports were not always considered so by experts. The share of non-scholarly journals ranged from 11% in experimental psychology to 25% in public administration.

Table 21.6. Scholarliness of journal articles by field: Burnhill and Tubby-Hille, UK social science

Field	*% of journal articles (468 total across all fields)*			*% of total publications*
	Peer–reviewed	*Authors consider scholarly*	*Other*	*Books*
Psychology	87	7	5	11
Statistics/computational methods	75	13	13	8
Geography & planning	73	19	8	7
Political science & internat. relations	64	8	28	29
Economics	64	6	30	10
Social anthropology	63	0	37	22
Management & business studies	60	12	29	10
Education	48	11	40	14
Sociology/ social administration	48	11	41	17
Economic & social history	44	20	37	24
Linguistics	23	15	62	20
All social science	62	13	26	15

If departmental output were recounted, including only articles in journals judged scholarly, in the best case one experimental psychology department would have lost only 1% of its output, and in the worst case one public administration department would have lost 61% of its output.

Nederhof et al. recalculated the share of articles covered by the SSCI in two ways based on their survey results. They calculated the share of articles in scholarly journals that were indexed in the SSCI, and they calculated the share of 'core' journal articles indexed in the SSCI where core journals were those:

1. known to more than 20% of their respondents;
2. possessing a high scholarly quality (mean of at least 7.5 on a 10 point scale);
3. and found useful to the research of at least 2 0% of the respondents.

Table 21.7 displays their results. The table shows that when just the scholarly core of a field is considered, SSCI coverage can be quite comprehensive. However, some fields remain mostly local in orientation. In public administration, a core literature could not even be identified.

Schoepflin (1990) reported similar results derived from a survey of German professors asked to rate journals according to their visibility and their perceived value. Of the highly rated journals the SSCI covered: 94% of psychology journals, 26% of sociology journals, and 8% of education journals.

Table 21.7. Share of articles indexed in SSCI by journal type — Dutch Social Science (% and number of articles)

Field	*University Annual Reports*	*Scholarly journals*	*Core journals*
Experimental psychology	58 (260)	69 (257)	100
General linguistics	21 (38)	22 (38)	85
Dutch language	10 (27)	11 (27)	20
Public Administration	3 (12)	5 (12)	no core

We can take two perspectives on this issue. In the first we ask: how good is the SSCI as a tool for evaluating Mode 1 social science? Clearly the value of the SSCI for evaluation increases when non-scholarly literature is removed from consideration. However, if we were to accept the mode II emphasis on knowledge in interaction with application, we would have to accept the importance of enlightenment literature. In recent years the culture of science has shifted to embrace the value of application and patenting. However, for social scientists this will be more difficult, in part because social science has always interacted with application and an internal tension has developed involving bolstering claims to scientific, and hence scholarly, status by distancing from application. Also, unlike the patent literature, the enlightenment literature has no review and citation mechanisms and so offers no differentiators by quality and extent of use, severely restricting the scope for assessment and evaluation.

6. CONCLUSION

In social science there are four distinct literatures: international journal articles, books, national and non-scholarly publications. International journal articles are SSCI indexed and are the currency of evaluation around the world. This is not wrong; using journal articles to communicate research results to an international audience is an important part of scholarly work. However, there is more to scholarly work in social science and the humanities. Books also can have a very high impact. National literature represents knowledge developed in a local context. Non-scholarly literature represents knowledge reaching out to application. To add to the problems each literature is more trans-disciplinary than comparable scientific literature. SSCI bibliometric evaluation must make the best of the low citation rates associated with trans-disciplinary citation scatter and citation accumulation times which are too long for policy makers' purposes. The authors and topics associated with the four literatures overlap, but not completely, so the results of SSCI bibliometrics will not be the same as the results of an ideal evaluation which included all four literatures.

All is not lost however, fields differ in their characteristics with the economics and psychology literatures quite similar to scientific literatures, sociology being a paradigmatic social science literature and history representing humanities. SSCI–based bibliometrics will work best when applied to science–like literatures such as economics and psychology.

Although scholarship around the world is moving into SSCI indexed journals, making standard bibliometrics more reasonable, the three other literatures still exist. If scholars seek to bolster their evaluations by abandoning the three other literatures in favour of SSCI journals, the resulting social science will differ from the social science of four literatures each serving specific ends.

REFERENCES

Bourke, P., Butler, L., Biglia, B. (1996). *Monitoring research in the Periphery: Australia and the ISI Indices*. Research Evaluation and Policy Project, Monograph Series No. 3, Canberra: Australian National University.

Broadus, R.N. (1971). The literature of the Social Sciences: a survey of citation studies. *International Social Sciences Journal*, 23, 236–243.

Burnhill, P.M., Tubby-Hille M.E. (1994). On measuring the relation between social science research activity and research publication. *Research Evaluation*, 4 (3), 130–152.

Butler, L. (1998). Personal communication of unpublished data.

Clemens, E.S., Powell, W.W., McIlwaine, K., Okamoto, D. (1995). Careers in print: books, journals, and scholarly reputations. *The American Journal of Sociology*, 101 (2), 433–494.

Cronin, B., Snyder, H., Atkins, H. (1997). Comparative citation rankings of authors in monographic and journal literature: a study of Sociology. *Journal of Documentation*, 53 (3), 263–273.

De Solla Price, D.J. (1970). *Citation measures of hard science, soft science, technology, and non-science*. In C.E. Nelson, D.K. Pollak (Eds.), Communication among Scientists and Engineers (pp. 1–12). Lexington, Mass: Heat.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M. (1994). *The new production of knowledge. The dynamics of science and research in contemporary societies*. London: Sage.

Glänzel W., Schoepflin, U. (1994). A stochastic model for the ageing analyses of scientific literature. *Scientometrics*, 30 (1), 49–64.

Glänzel W., Schoepflin, U. (1995). A bibliometric study on ageing and reception processes of scientific literature. *Journal of Information Science*, 21 (1), 37–53.

Glänzel, W. (1996). A bibliometric approach to social sciences. National research performances in six selected social science areas, 1990–1992. *Scientometrics*, 35 (3), 291–307.

Glänzel W., Schoepflin, U. (1999). A bibliometric study of reference literature in the sciences and social sciences. *Information Processing and Management*, 35, 31-44.

Glänzel W., Schubert, A., Schoepflin, U., Czerwon, H.J. (1999). An item–by–item subject classification of papers published in journals covered by the SSCI database using reference analysis. *Scientometrics*, 46 (3), 431–441.

Godin, B. (2002). *The social sciences in Canada: what can we learn from bibliometrics*? INRS, Working Paper no 1. Quebec, Canada: INRS.

Hicks, D., Potter, J. (1991). Sociology of scientific knowledge: a reflexive citation analysis or science disciplines and disciplining science. *Social Studies of Science*, 21, 459–501.

Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics*, 44 (2), 193–215.

Hicks, D. Breitzman, A., Olivastro, D., Hamilton, K. (2001). The changing composition of innovative activity in the U.S.—a portrait based on patent analysis. *Research Policy*, 30, 681–703.

Ingwersen, P. (1997). *The central international visibility of Danish and Scandinavian research 1988–1996.* A general overview of science & technology, the humanities and social sciences by online publication analysis., 17 p. (CIS Report 5.3).

Ingwersen, P., Wormell, I. (1999). Publication behaviour and international impact: Scandinavian clinical and social medicine 1988–96. *Scientometrics*, 46 (3), 487–499.

Ingwersen, P. (2000). The international visibility and citation impact of Scandinavian research articles in selected social science fields: the decay of a myth. *Scientometrics*, 49, 39–61.

Ingwersen, P. (2002). Visibility and impact of research in psychiatry for North European countries in EU, US and world contexts. *Scientometrics*, 54, 131–144.

Katz, J.S. (1999). *Bibliometric indicators and the social sciences*. Report prepared for UK economic and social research council, www.sussex.ac.uk/Users/sylvank/pubs/ESRC.pdf

Kyvik, S. (1988). Internationality of the social sciences: the Norwegian case. *International Social Science Journal*, 163–172.

Lewison, G. (2001). Evaluation of books as research outputs in history of medicine. *Research Evaluation*, 10 (2), 89–95.

Leydesdorff, L. (2003). Can networks of journal–journal citation be used as indicators of change in the social sciences? *Journal of Documentation*, 59 (1), 84–104.

Line, M.B. (1979). The influence of the type of sources used on the results of citation analyses. *Journal of Documentation*, 35 (4), 265–284.

Luwel, M., Moed, H.F. Nederhof, A.J. De Samblanx, V. Verbrugghen, K., Van Der Wurff, L.J. (1999). *Towards indicators of research performance in the social sciences and humanities*. An exploratory study in the fields of Law and Linguistics at Flemish Universities. Report of the Flemish Inter-University Council (V.L.I.R.), Brussels, Belgium / Centre for Science and Technology Studies (CWTS), Leiden University, the Netherlands / Ministry of the Flemish Community, Brussels, Belgium. V.L.I.R.: Brussels, Belgium.

Moed, H.F., Nederhof, A.J., Luwel, M. (2002). Towards performance in the humanities. *Library Trends* (Special Issue on current theory in library and information science). 50, 498–520.

Narin, F. (1994). Patent Bibliometrics. *Scientometrics*, 30 (1), 147–155.

Narin, F., Hamilton, K.S., Olivastro, D. (1997). The increasing linkage between U.S. technology and public science. *Research Policy*, 26 (3), 317–330.

Nederhof, A.J. (1989). Books and chapters are not to be neglected in measuring rsearch productivity. *American Psychologist*, 44, 734–735.

Nederhof, A.J., Zwaan, R.A. DeBruin, R.E. Dekker, P.J. (1989). Assessing the usefulness of bibliometric indicators for the humanities and the social and behavioural sciences: a comparative study. *Scientometrics*, 15 (5–6), 423–435.

Nederhof, A.J., Zwaan, R.A. (1991). Quality judgments of journals as indicators of research performance in the humanities and the social and behavioral sciences. *Journal of the American Society for Information Science*, 42 (5), 332–340.

Nederhof, A.J., Meijer, R.F. Moed, H.F., Van Raan, A.F.J. (1993). Research performance indicators for university departments: a study of an agricultural university. *Scientometrics*, 27 (2), 157–178.

Nederhof, A.J., Van Raan, A.F.J. (1993). A bibliometric analysis of six economics research groups: a comparison with peer review. *Research Policy*, 22, 353–368.

Nederhof, A.J., Van Wijk, E. (1997). Mapping the social and behavioral sciences world-wide: use of maps in portfolio analysis of national research efforts. *Scientometrics*, 40 (2), 237–276.

Nederhof, A.J., Van Wijk, E. (1999). Profiling institutes: identifying high research performance and social relevance in the social and behavioral sciences. *Scientometrics*, 44 (3), 487–506.

Nederhof, A.J., Luwel, M., Moed, H.F. (2001). Assessing the quality of scholarly journals in linguistics: An alternative to citation–based journal impact factors. *Scientometrics* 51 (1), 241–265.

Nowotny, H., Scott, P., Gibbons, M. (2001). *Re-thinking Science*. Cambridge, UK: Polity Press.

Nowotny, H., Scott, P., Gibbons, M. (2003). *Re-thinking science: mode 2 in societal context*. http://www.nowotny.ethz.ch/pdf/Nowotny_Gibbons_Scott_Mode2.pdf.

Pestaña, A., Gómez, I., Fernández, M.T., Zulueta, M.A., Méndez A. (1995). *Scientometric evaluation of R&D activities in medium–size institutions: a case study based on the Spanish Scientific Research Council (CSIC)*. In M. Koenig, A. Bookstein (Eds.), The Proceedings of the Fifth International Conference of the International Society for Scientometrics and Informetrics (pp. 425–434).

Pierce, S. (1987). Characteristics of professional knowledge structures: some theoretical implications of citation studies. *Library and Information Science Review* (LISR), 9, 143–171.

Royle, P., Over, R. (1994). The use of bibliometric indicators to measure the research productivity of Australian Academics. *Australian Academic & Research Libraries*, 25 (2), 77–88.

Schoepflin, U. (1990). *Problems of representativity in the Social Sciences Citation Index*. In P. Weingart, R. Sehringer, M. Winterhager (Eds.), Representations of science and technology, Proceedings of the International Conference on Science and Technology Indicators, Bielefeld, Germany, 10–12 June, 1992 (pp. 177–188). Leiden: DSWO Press.

Small, H., Crane, D., (1979). Specialties and disciplines in science and social science: an examination of their structure using citation indexes. *Scientometrics* 1 (5–6), 445–61.

Thomas, P. (1998). *A bibliometric analysis of fashions in management literature*, PhD thesis, Nottingham Trent University.

Thorsteinsdottir, H. (1998). *Islands Reaching Out*, unpublished DPhil thesis, University of Sussex.

Tijssen, R.J.W., Van Leeuwen, Th.N., Verspagen, B., Slabbers, M. (1996*). Wetenschaps- en Technologie–Indicatoren* 1996, Het Nederlands Observatorium van Wetenschap en Technologie: Centrum voor Wetenschaps- en Technologie Studies (CWTS) en Maastricht Economic Research Institute on Innovation and Technology (MERIT) in opdracht van het Ministerie van Onderwijs, Cultuur en Wetenschappen, Zoetermeer, (ISBN 90-75023-03-0), 223p.

Van Der Meulen, B., Leydesdorff, L. (1991). Has the study of philosophy at Dutch universities changed under economic and political pressures? *Science, Technology, & Human Values*, 16 (3), 288–321.

Villagrá Rubio, A., (1992). Scientific production of Spanish universities in the fields of social sciences and language. *Scientometrics*, 24 (1), 3–19.

Webster, B.M. (1998). Polish sociology citation index as an example of usage of national citation indexes in scientometric analysis of social science. *Journal of Information Science*, 24 (1), 19–32.

Winclawska, B.M. (1996). Polish sociology citation index (principles for creation and the first results). *Scientometrics*, 35 (3), 387–391.

Winterhager, M. (1994). *Bibliometrische Basisdaten zur Entwicklung der Sozialwissenschaften in Deutschland.* in H. Best, et al. (Hrsg.), Informations- und Wissensverarbeitung in den Sozialwissenschaften. Opladen 1994, 539–551.

Yitzhaki, M. (1998). The language preference in sociology. *Scientometrics*, 41 (1–2), 243–254.

Zwaan, R.A., Nederhof, A.J., (1990). Some aspects of scholarly communication in linguistics: An empirical study. *Language*, 66, 523–527.

APPENDIX

Data sources and method for Table 21.5

Study	*Data source*
Bourke et al. 96	IAS95 — Database of research output 1989 to 1993 for the Research School of Social Sciences (RSSS) and Research School of Pacific (and Asian) Studies (RSPAS), Institute of Advanced Studies (IAS), at the Australian National University (ANU).
Burnhill	All publications related to research grants of the UK Economic and Social Research Council (ESRC) in 1984–85.
Butler	IAS95 database though with some non-ANU university papers included, figures from personal communication
Hicks & Potter 91	Bibliography of sociology of scientific knowledge collected by snowball method.
Nederhof, 89	Bibliographies of Dutch university output in eight fields constructed by correcting lists obtained from university annual reports. All figures averaged across the eight fields.
Nederhof et al., 93	Grant related bibliographies of six British economics research groups
Pestaña	Bibliography constructed from Annual Reports of the Spanish Scientific Research Council (CSIC) 1990–92
Royle & Over	Bibliography of articles published in journals or serials constructed from the 1990 and 1991 Annual Reports of La Trobe University, Monash University, and the University of Melbourne
Tijssen	Research papers of Dutch universities, personal communication
Villagra Rubio 92	Database of Spanish university journal and book output in economics, sociology, political science, linguistics and literary sciences derived from the ECOSOC database which contains all articles in Spanish journals and the ISBN database, the official bibliography of Spanish books. These were supplemented with searches in 11 international databases including Social Scisearch.
Winterhager	Das Sozialwissenschaftliche Literaturinformationssystem (SOLIS)

Chapter 22

EVALUATION OF RESEARCH PERFORMANCE AND SCIENTOMETRIC INDICATORS IN CHINA

Bihui Jin[1] and Ronald Rousseau[2,3]

[1]*Documentation and Information Centre of the Chinese Academy of Sciences, Chinese Science Citation Database,Beijing, P.R. China*
[2]*Antwerp University, Information and Library Science (IBW), Wilrijk, Belgium*
[3]*KHBO, Department of Industrial Science & Technology, Oostende, Belgium*
E-mail: ronald.rousseau@khbo.be

Abstract: Chinese scientists and decision makers are, like their counterparts elsewhere in the world, highly interested in analysing the quality of their country's academic and technological achievements. Twenty five years ago research activities Chinese scientists understood the world's scientific research largely through journals consulted in libraries. in China were mainly conducted within the nation's borders. Chinese scientists understood the world's scientific research largely through journals consulted in libraries. They themselves published few papers in international journals. Meanwhile, the number of Chinese publications, as covered by SCI, has increased spectacularly: it rose from about 8,000 in the year 1990 to nearly 40,000 in the year 2002. What is the explanation for this extraordinary growth? The fundamentals of this accomplishment are to be found in the country's major socio-economic development and the stimulating role of the government's S&T policy. Although the number of Chinese publications covered by the *Web of Science* has rapidly increased, most Chinese research results are still published in domestic journals. This means that it makes little sense to use the Web of Knowledge as the only source for information retrieval, or for research evaluation purposes in China. For this reason it had already been decided in 1989 to develop local, i.e., Chinese, citation databases. This was the origin of the *Chinese Science Citation Database* (CSCD) and the *China Scientific and Technical Papers and Citations* (CSTPC). In 1998 a *Chinese Social Sciences Citation Index* (CSSCI) was developed as well. Local databases of derived indicators, similar to the Journal Citation Reports, soon followed. The structure of these databases is described here. This contribution focuses on two parallel developments: one using ISI's databases, aimed at gauging China's international position; and one used for internal purposes, where the locally developed databases play an important role. Examples of comparisons and

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 497-514.

rankings based on local databases are given, and it is shown how these are different from rankings based on ISI's databases. Chinese scientists and decision makers soon recognized that simple quantitative evaluations focussing on numbers stimulate the growth of publications, but have little effect on the quality of research. Hence new approaches and regulations for research evaluation are nowadays being introduced.

1. INTRODUCTION

Since 1989 China has paid special attention to scientometric indicators assessing its scientific position in the world. Indeed in that year the Institute of Scientific and Technical Information of China (ISTIC) announced for the first time ISI's new statistical results through the news media. Over the years publication numbers of the main scientific countries and their ranking have more and more attracted the attention of Chinese scientists and decision makers. It is realised that such quantitative data can help understanding and analysing the position and development of science in China. Insight in these matters leads to better strategic planning and a more targeted S&T policy.

Owing to historic and linguistic reasons Chinese scientists traditionally publish most of their research results in Chinese and in domestic journals. This implies that it shows poor judgement to use ISI's databases as the only source for information retrieval, or for research evaluation purposes in China. For this reason it had already been decided 1989 to develop local, i.e., Chinese, citation databases. This was the origin of the Chinese Science Citation Database (CSCD) and the China Scientific and Technical Papers and Citations (CSTPC) (Meng & Wang, 1996; Jin & Wang, 1999; Group for Statistics and Analysis of Chinese Articles, 2001; Wu et al., 2003). In 1998, a Chinese Social Sciences Citation Index (CSSCI) was developed as well (Su et al., 2001). Local databases of derived indicators, similar to ISI's Journal Citation Reports, followed soon (Jin et al., 2002). In combination with ISI's databases these play an important role in conducting research evaluation and in the quantitative study of science and technology in China.

Nowadays, after more than ten years of experience, a discussion about the role of quantitative data and methods of research evaluation is taking place in China's scientific community. Chinese scientists and decision makers are rethinking the problem and drawing lessons from past practice. These questions are discussed further on in Section 4.3.

2. THE CONSTRUCTION OF LOCAL SCIENCE CITATION DATABASES

According to data issued by the Ministry of Science & Technology there were 4835 research institutions in China in 1996 (China Science & Technology Monthly, 1998) while 2,071,530 scientists and engineers were active in 2001 (National Bureau of Statistics & Ministry of Science and Technology, 2002). As stated by another statistical source, 4,420 titles of science and technology journals were published in 2001 (Association of Chinese Publishers, 2002). It is therefore estimated that China annually produces about half a million scientific papers.

In order to improve China's knowledge flow, including searching and retrieving scientific information, and optimally evaluate scientific and technological performance, it is absolutely necessary to optimise the use of international bibliographic, citation, and indicator databases, and exchange experiences in their utilisation, making use of feedback obtained from senior scientists and institutional directors, and at the same time, develop and improve China's own citation databases, reflecting the characteristics of domestic activities.

2.1 An Overview of the Chinese Science Citation Database

With the actual demands of Chinese scientists and policy makers in mind, the Documentation & Information Centre of the Chinese Academy of Sciences (DICCAS) started in the late 1980s a small scale experiment on the feasibility of CSCD. In 1991 CSCD became a research and development project co-supported by the National Science Foundation of China (NSFC) and the Chinese Academy of Sciences (CAS). After a period of almost fifteen years of construction and efforts, the CSCD has evolved into a large database with multiple integrated functions. As such, it has earned a good reputation amongst Chinese scientists.

In its early stage the CSCD covered only 315 journals published in China. In 1996 the number of CSCD source journals was expanded to 582 titles, growing to 1,046 titles in 2002. This accounts for approximately twenty-three percent of the total number of science and technology journals of China. Source journals of the CSCD are journals with an emphasis on basic research and theory, or research in high tech areas. Indicators for the evaluation of journals (when screening for possible inclusion in the CSCD) are: the journal's overall citation rate; its impact factor; its coverage by international databases; and its inclusion in the *Chinese Core Journal List* compiled by Peking University. At present the CSCD covers almost all of the science and technology fields, including mathematics, physics, chemistry, astronomy, geosciences, biology, agriculture and forestry sciences, medical sciences, engineering and technology, environmental sciences, and management science. Between 1989 and 2001 the CSCD accumulated about 860,000 source records (articles), leading to 2.4 million citation records (references). Note that the CSCD only covers citations of Chinese articles (cited in Chinese journals, but not necessarily published in Chinese journals).

Besides a number of basic bibliographic fields and functions, similar to those of the SCI, the CSCD has developed some special fields, making it a richer source of information. These include publication lags (time between acceptance and publication), the funding agency, and the age and gender of authors. Based upon the mature and advanced technology of the SCI, and by adding a number of new functionalities, the CSCD has become a locally built authoritative citation retrieval tool for China.

2.2 Scientometric Indicators Based on the SCI and the CSCD

The CSCD is an important complement to ISI's databases for China. The database is, however, not yet publicly available through vendors such as DIALOG. Plans have been made for this, and in the near future the CSCD will be available through the Internet. Based on a combination of the resources offered by the CSCD and the SCI, DICCAS designed in 1998 a set of Chinese Scientometric Indicators (CSI) for research evaluation. These indicators are published in two versions: a printed version, under the name of Chinese Scientometric Indicators (CSI), and a CD-ROM version, under the name of Chinese Scientometric Indicators Database (CSID). The contents of these two versions are the same, but, of course, the CD-ROM version offers more functionality and greater flexibility in use. More details about this derived database can be found in (Jin et al., 2002).

CSI has nearly 200 indicators divided over eight subjects, discussed below in sections 2.2.1 to 2.2.8.

2.2.1 Statistical sources

It should be noted that since 1999 the CSCD has been divided into two parts: a core part, the source of the basic statistics provided by the CSI and the expanded part, used only for search and information retrieval. In 2002 the two parts together covered 1,046 journals of which 670 form the core.

This section contains statistics on the subject distribution of Chinese articles in the SCI and CSCD source journals, and on the distribution of SCI source journals by country and region. Data are given on the ranking zones of journals containing articles by Chinese scientists (see further for more details about the method of determining these zones). It further contains the publication distribution of articles in different subjects (for the SCI and the CSCD) and the citation distribution in the CSCD of articles in different subjects. This subset of indicators is the basis for all other ones. Rankings differ according to the source used, showing the complementary nature of the two databases. An example is given in Table 22.1.

Table 22.1. Different rankings for top 10 universities of China in terms of publications according to SCI and CSCD in 2001

Top 10 Universities as ranked by SCI	*Articles in SCI*	*Rank by CSCD*	*Top 10 Universities as ranked by CSCD*	*Articles In CSCD*	*Rank by SCI*
1. Tsing Hua Univ	1,024	2	1. Zhejiang Univ	2,706	5
2. Beijing Univ	788	3	2. Tsing Hua Univ	2,635	1
3. Nanjing Univ	702	20	3. Beijing Univ	2,443	2
4. S&T Univ China	602	13	4. Huazhong Univ S&T	2,021	15
5. Zhejiang Univ	526	1	5. Shanghai Jiao Tong Univ	1,780	8
6. Fudan Univ	453	8	6. Sichuan Univ	1,556	13
7. Shandong Univ	366	25	7. Fourth Mil Med Univ	1,521	48
8. Shanghai Jiao Tong Univ	361	5	8. Fudan Univ	1,432	6
9. Jilin Univ	326	15	9. Xian Jiaotong Univ	1,358	21
10. Nankai Univ	304	34	10.Beijing Union Med Coll	1,348	17

2.2.2 Institutional indicators

The number of papers and citations are given for universities and colleges, research institutes, and medical institutions. Following a Bradford –like approach, the SCI and the CSCD journals were subdivided respectively into four and three zones. A unified method for assessing international (SCI) and domestic (CSCD) articles was proposed in order to determine the publication productivity of an institution (Jin et al., 1999). Publication numbers and impact (number of citations) were obtained for all Chinese institutes. Not surprisingly, these numbers are heavily skewed. A distinction has been made between the most active ones, responsible for 30% of all publications and citations and the other ones. The most active ones are referred to as core institutes, and only their data are published in the CSI.

2.2.3 State key laboratory indicators

There are 159 state key laboratories and more than one hundred open laboratories in China. The Chinese government gives these laboratories special support. In return these labs must accept government evaluation. Responding to this situation, the CSI provides a set of indicators for state key laboratories and open laboratories. The number of articles of these labs, number of citations, number of authors, subject distribution of articles, and the number of articles supported by the NSFC have been counted and tabulated, see Table 22.2 for an example.

Table 22.2. Number of publications of state key labs and open labs in 2001

Labs	*SCI Papers*	*CSCD papers*
State Key Labs	2,820	5,701
Open Labs of Ministries	1,356	4,058

2.2.4 S&T funds indicators

Setting up a system for the funding of science is one of the important policy measures taken during the recent research reform of China. A funding system cascading from the central government to local governments was set in place. In this system the NSFC plays a major role for the funding of basic research.

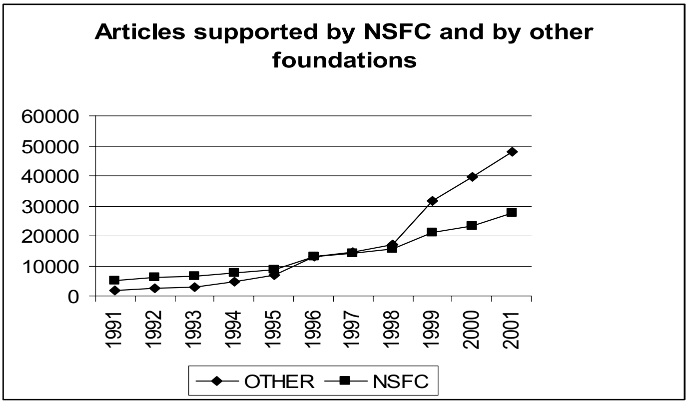


Figure 22.1. Number of articles produced with support of scientific foundations

Corresponding to its role in research evaluation the CSI also procures statistics on the number of articles published with the help of these funds and gives a ranking of colleges and research institutions with respect to the number of articles supported by the NSFC (see Figure 22.1).

2.2.5 Regional indicators

The People's Republic of China is subdivided into 31 provinces and autonomous regions. These regions are characterised by an unbalanced development. This is true in the areas of economy and culture as well as S&T (Tsui, 1996; Yang, 1999; Démurger, 2001; Jin and Rousseau, 2001). The CSI provides statistics on the number of articles and the subject distribution for each province and autonomous region. The number of core institutions of each province and autonomous region is presented based on the data from the institution indicator subset. These data can be used to analyse the disciplinary advantages and weaknesses of each province and autonomous region. Not surprisingly, Beijing and Shanghai are the most productive regions in China, followed by the provinces of Jiangsu and Guangdong. More details about regional inequality in publication output can be found in (Jin and Rousseau, 2001).

2.2.6 Indicators of research collaboration

Collaborative research is a general trend in modern science. The number of articles written in collaboration with other nations, those resulting from collaboration among provinces and autonomous regions and, finally, articles written in a collaborative effort among universities, institutes and enterprises have all been counted and tabulated.

2.2.7 Author indicators

The distribution of gender, age, academic degree (MSc. or Ph.D.) for Chinese authors based on data obtained from the original publications (in about 50% of the cases this information is provided in Chinese journals) are collected. These data are useful for analysing and evaluating the sociological structure of Chinese research personnel (Jin et al., 2003).

2.2.8 Literature indicators

Citation analysis is an important tool in journal evaluation. For a worldwide assessment of journals the JCR contain the best (or most often used) statistics. Similar statistics on the cited frequency of Chinese journals and their (local) impact factors have been obtained from the CSCD. Journals with the highest impact factors in different disciplines are shown in Table 22.3. Note that all of these journals publish in Chinese. Some of them have an English edition, covered by ISI. The English edition of the Chinese Science Bulletin, for instance, is covered by ISI, with a JCR–impact factor of 0.511 (2001).

Table 22.3. Journals with highest CSCD impact factor in each subject category (2001)

Subjects	*Journals (translated titles)*	*Impact factor CSCD*
MATH	J SYST ENGINEERING	0.373
PHYS	ACTA PHYSICA SIN	0.867
CHEM	CHEM J CHINESE UNIV	0.842
GEOSCI	ACTA GEOLOG SIN	1.934
BIOL	ACTA BIOCH BIOPH SIN	0.814
MED	CHINESE J VIROL	0.566
AGRI	ACTA PEDOLOGICA SIN	0.664
ENGIN	CHINESE J NONFERROUS MET	0.740
ENVIRON SCI	ACTA ECOL SIN	0.871
MANAGEMENT SCI	J MANAG SCI CHINA	0.491
MULTI-DISC	CHINESE SCI BULL	0.665

3. A UNIFIED METHOD OF WEIGHTING INTERNATIONAL AND DOMESTIC ARTICLES

Combining data from two different databases, one taking an international perspective and the other taking a domestic point of view, is not straightforward. Yet policy makers wanted to see a global ranking and weighting based on the combination of the two data sets. As a result we proposed the following ad hoc unified weighting method of counting publications (Jin et al., 1999). We are fully aware that this is at best a list of impact potential, not necessary correlated with the impact actually obtained.

In an attempt to compare, as much as possible, like with like, journals have first been distributed over twelve subject categories: mathematics; physics; chemistry; astronomy; geosciences; biology; agriculture and forestry sciences; medical sciences; engineering and technology; environmental sciences; management science; and a multi-disciplinary category. We denote the number of journals in a subject category by n_i, $1 \leq i \leq 12$.

The unified method has been applied to each category separately and consists of two steps. In the first step journals have been assigned to zones. The second step assigns a weight to each zone.

3.1 Zones in a Category of ISI Journals

First, each journal in the category has been assigned an average impact factor (AIF), namely the average of the latest three standard, i.e., ISI two year impact factors. In the year 2003, for example, this average impact factor is:

$AIF = (IF_{2001} + IF_{2002} + IF_{2003})/3$

Journals have then been ranked according to the AIF. The AIF of the j^{th} ranked journal of category i is denoted as AIF_{ij}.

The first zone always consists of the top *5%*. We denote the number of journals in the first zone of category i by m_{i1}.

The second zone of category i consists of the journals ranked $m_{i1}+1, \ldots, m_{i2}$, so that

$\sum_{j=m_{i1}+1}^{m_{i2}} AIF_{ij} \Big/ \sum_{j=m_{i1}+1}^{n_i} AIF_{ij} = 1/3$. Similarly, the third zone consists of the journals ranked $m_{i2}+1, \ldots, m_{i3}$ such that $\sum_{j=m_{i2}+1}^{m_{i3}} AIF_{ij} \Big/ \sum_{j=m_{i1}+1}^{n_i} AIF_{ij} = 1/3$.

Finally, the fourth zone consists of the remaining journals, i.e. those ranked from $m_{i3}+1$ to n_i. Their AIFs satisfy the relation: $\sum_{j=m_{i3}+1}^{n_i} AIF_{ij} \Big/ \sum_{j=m_{i1}+1}^{n_i} AIF_{ij} = 1/3$.

3.2 Zones in a Category of CSCD Journals

Also for these journals the AIF is used, but now based on impact factors calculated from the CSCD. Once the journals have been ranked based on AIF values, the whole list has been subdivided into three zones in exactly the same way as has been done for the remaining *95%* ISI journals.

We have now seven zones of journals. The next step is to assign a weight to each of these zones.

3.3 Assigning a Weight to Each Zone

The average AIF value of each zone is determined, and each zone receives a weight equal to the quotient of this average, and the average of the third CSCD zone. Hence the third CSCD zone receives a weight equal to one. The weight of a journal is then the weight of the zone this journal belongs to.

As the pool of journals from which an article (and hence a journal) may receive citations is much larger for ISI journals than for CSCD journals, it is quite natural that, in general, CSCD impacts are smaller than ISI journals. As it happened that the weight of the fourth ISI zone was usually close to the weight of the first CSCD zone, it was decided to join these two zones into one (with an average weight). Moreover Chinese journals covered by ISI (and hence also covered by the CSCD) usually belong to this merged zone. Such journals receive this weight twice (promoting Chinese journals covered by ISI). Note though that the weight of the third zone is more than twice the weight of the fourth. The following figure (Figure 22.2) illustrates the procedure.

JCR and CSCD journals	*JCR journals*	*CSCD journals*
First zone: 289	First zone: 289	
Second zone: 694	Second zone :694	
Third zone: 1,268	Third zone: 1,268	
Fourth zone: 3,597	Fourth zone: 3,497	First zone: 100
Fifth zone: 174		Second zone: 174
Sixth zone: 369		Third zone: 369

Figure 22.2. Connecting four zones of JCR with three zones of CSCD in 2001

4. EVALUATION OF RESEARCH PERFORMANCE IN CHINA

About twenty years ago China did not perform any quantitative evaluation of research. Since the reform of the scientific system, however, quantitative evaluation has been introduced into research management and decision making related to S&T. In recent times the attitude with respect to quantitative methods has greatly changed. Administrators responsible for funding, journal editors, scientists, engineers, and every one involved in the scientific enterprise are all coming to terms with bibliometric and scientometric methods. From a mild interest, quantitative methods have turned into an almost daily practice.

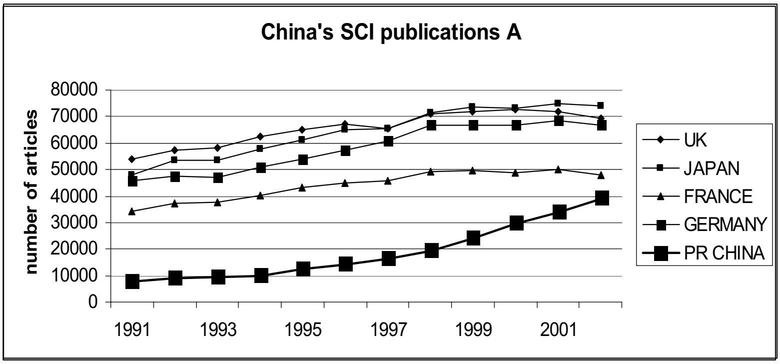
4.1 Remarkable Increase of Chinese Publications

Before the year 1989 few Chinese scientists knew of the Institute of Scientific Information (ISI) and its products. The majority published their papers in whatever journals they preferred. In the year 1989 the Institute of

Scientific and Technical Information of China (ISTIC) introduced the bibliometric method for the evaluation of research performance to the larger scientific community. Rankings of countries, provinces, institutions, and scientists based on the data of databases such as SCI, Engineering Index (EI), Index to Scientific and Technical proceedings (ISTP) appeared first in public. Since then ISTIC has yearly announced SCI's new statistical results through the news media. Now more and more Chinese scientists and decision makers are paying attention to ISI's lists of rankings, and are choosing the journals covered by ISI to publish their papers.

Figure 22.3 shows that the number of Chinese publications has strongly increased since the year 1991. According to SCI data China published only 3475 articles in 1983, increasing to 7705 in the year 1991. In the year 2002 this figure has further risen to 39013. This is a truly exponential growth: indicating the year 1991 by 1 and hence the year 2002 by 12 yields the relation: *NUMBER of PUBLICATIONS = 5492*EXP(0.164*YEAR)*, with an R^2 value of 0.99 (nonlinear regression). The ranking of publications of China in the world, according to the SCI, rose from the 24th place in 1985, to the 14th in 1991, and further to the 6th in 2002. During this period China has overtaken countries such as Belgium, the Netherlands, India, Spain, Australia and Canada in publication output. It is now only behind USA, UK, Germany, Japan and France. Data for the period before 1991 have been taken from (Shang, 1988).

What are the reasons for this increase? Two types of reasons may be distinguished: those external to the scientific system, and internal reasons. As external reasons we mention the return of Hong Kong to China in 1997. In the year 2002 the former Crown colony on its own was responsible for 5,827 SCI papers. Another external reason is the fact that during the last decade the Chinese national economy is among the fastest developing ones in the world. This leads to an environment favourable to research, encouraging investments in R&D. China is not just a country of cheap labour and of production of low cost value added goods, but more and more a region where international firms are competing based on capital – and technology–intensive strategies (Li et al., 2000). As internal reasons we mention that Chinese research evaluation policies strongly support publications in journals covered by ISI. Another reason internal to the science system is that more and more Chinese journals are being covered by the SCI. In the year 1991 only 14 Chinese journals and 1,910 papers were covered by the SCI. In the year 2002 there were already 68 Chinese journals publishing 11,399 papers included in the SCI –Expanded. Moreover, a fast increase of papers with Chinese authors in international journals is clearly noticeable.



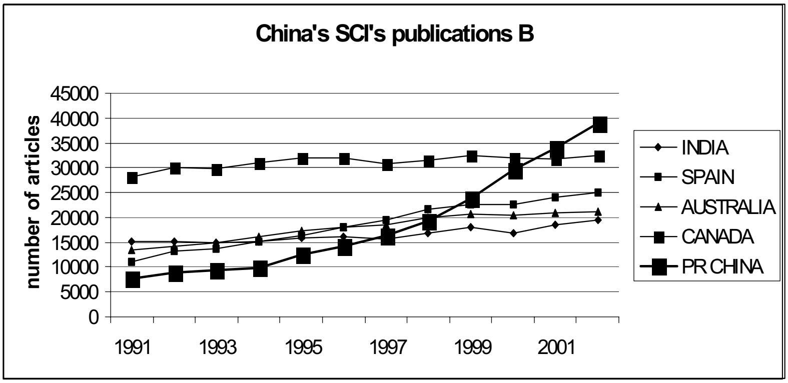


Figure 22.3a, b. Evolution of the number of Chinese articles published in journals covered by the SCI. Comparison with major countries

4.2 The Number of Papers with International Collaboration Increases Quickly

International collaboration is an important strategy for reducing the S&T lag between China and Western countries. Over the last twenty years a period characterized by reforms and openness to the West, international collaboration in China's research system has developed rapidly. This can be concluded from the increasing number of publications resulting from international collaborations including Chinese scientists. This is not only visible in international journals, but also, though to a smaller extent, in journals published in China. In this respect we note the following three characteristics: publications with international collaboration exhibit an increasing upward trend; the bulk of these publications appear in international journals covered by ISI; and finally, although collaborations take place with more than one hundred countries, they mainly concentrate on a few preferential partner countries, such as the USA, Japan, Germany and, increasingly, Australia (Wang & Wu, 2001). Figures 22.4 and 22.5 illustrate these trends.

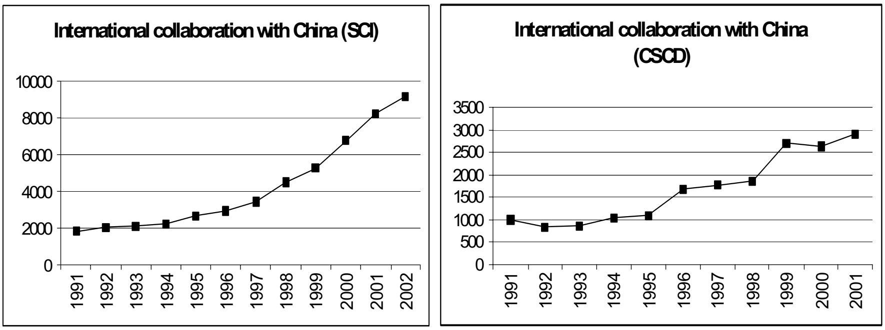


Figure 22.4. International collaboration with China (data based on the SCI and on the CSCD)

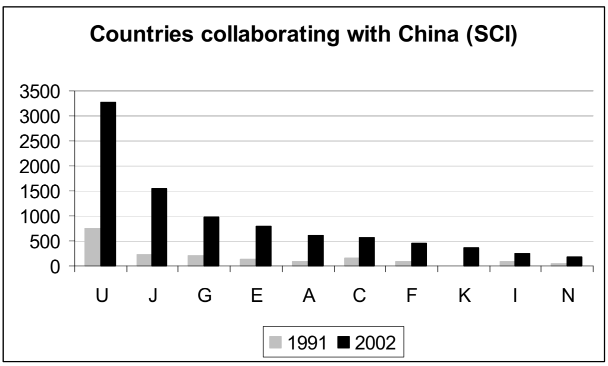


Figure 22.5. Evolution of collaboration with China: preferred countries. U = USA; J = Japan; G = Germany; E = England. A = Australia, C = Canada; F = France; K = South Korea. I = Italy; N = Netherlands

4.3 The Practice of Quantitative Evaluation in China

The practice of quantitative evaluation in China began in the early nineties. Before that time research evaluation was done by peer review or executive decision–making. Peer review and executive decision–making led to many problems as the result of biased decisions, favouring 'famous' people, and established institutes and universities. Consequently, this approach ran into a credibility crisis. Against this background some research management departments introduced the method of quantitative evaluation.

In the second half of the nineties the CSCD and CSTPC, fine-tuned to the Chinese situation, came more and more into use (Wu et al., 2003). From that time on the method of quantitative evaluation has rapidly gained acceptance in research institutions and management departments.

In practice many indicators are involved in quantitative evaluation. Some are purely numerical data, such as the number of papers published, or the number of citations received. Other ones are relative measures, such as the number of citations received per paper, as compared to the average in the field (Russell & Rousseau, 2002).

Methods used in evaluation procedures influence the way in which science is done: this inevitably leads to wanted and unwanted, positive and negative effects. This is also true for the Chinese practice of quantitative evaluation. The exponential increase of Chinese publications in the SCI is one of the positive effects. This in turn is leading to a strong positive effect on the international visibility of Chinese science. More and more Western scientists are looking forward to collaborating with Chinese colleagues, appreciating their specific ‘know-how’. On the other hand, all too often long-term research quality is being ignored. Many researchers are concentrating on ‘fashionable’ research, preferably of the kind that leads to fast results. Publishing in ISI–journals is becoming the highest academic standard, instead of real research quality. In some situations quantitative data are playing a decisive role in performance evaluation, because decision makers do not have enough time, or the expertise, to understand the details of research results. Such are the unwanted consequences of introducing (only) quantitative measures in research management, leading to new problems faced by decision makers. Of course, the same phenomenon is going on in the West. One could say that Chinese scientists have been ‘learning’ from the West also in this respect.

The government has recently issued two important documents: *Decision on Improving S&T Evaluation* (Chinese Basic Science, 2003) and *Methods and Techniques of S&T Evaluation* (Science & Technology Ministry of the P.R. China, 2003). These documents make a strong case for an evaluation based on quality considerations, emphasise the role of peer review as a base for improving the present system and regulations, and call for a rational use of quantitative data and scientometric indicators.

5. CONCLUSION

5.1 Sharp Contrast in Numbers of Publications and Citations

According to recent Essential Science Indicators developed by ISI, China ranks 9th by publications and 19th by citations over the period 1993–2002. These ranks bear testimony to the progress China has made over the past ten years. Yet Chinese citations are still at a relatively low level. The numbers are not proportionate to the number of Chinese publications. It is even true that publishing a Chinese article decreases the impact factor of a journal (the so-called Matthew effect for countries and journals) (Bonitz et al., 1999). Research getting high citation scores is often work at the frontier of science. Low citation counts point to Chinese science still being at the periphery of world research. This statement should, however, be qualified. Citations only reflect the past, and are subject to time lags. Moreover, China has launched several ambitious plans to join the world scientific elite (Jiang, 2000; Cyranoski, 2001).

5.2 The Special Role of Local Science Citation Databases in non-English Speaking Countries

The *Science Citation Index* initiated by Eugene Garfield is a unique retrieval and evaluation tool. Yet it is known that it is not adequate for the local evaluation of less developed non-English speaking countries, or for the retrieval of these countries' publications. The SCI is a good tool for China to watch the world (of science) and for the world to watch China (Moed, 2002). If, however, one is interested in Chinese S&T as a whole, the SCI has not enough authority. Even Chinese journals included in the ISI do not enjoy a large visibility (Ren et al., 1999; Ren and Rousseau, 2002). For this purpose it is necessary to combine the SCI with local citation databases. Such citation databases are complementary to ISI's (Liang et al., 2001; Liang, 2003).

5.3 Limitations of Bibliometric Indicators at the Micro Level, as Perceived in China

What roles do scientometric and bibliometric indicators play in research evaluation? Over recent years there is an extensive discussion going on about this problem in China (Wang, 2001; Wu and Liang, 2001; Sun and Xu, 2002; Ren, 2002; Wu, 2002). From the practice of the past decade Chinese scientists and decision makers have understood that, at least at the macro level, analyses based on scientometric and bibliometric indicators may provide a benchmark and bring trends to the fore. They help decision – makers in gauging China's position in the world. At the micro level, in particular in performance evaluations of scientists, scientometric and bibliometric indicators have severe limitations. For really ground breaking research the required activities are complex phenomena performed by human beings with high intelligence. Even for such gifted scientists, planning as well as perseverance and serendipity play essential roles. At the individual level research and research results are not continuous phenomena: it proceeds with ups and downs. Quantitative data can not be used for this. Qualitative evaluation by impartial peer review should play a main role at this level. Even at higher aggregation levels quantitative evaluation of research performance and bibliometric indicators are not intended to replace qualitative peer review, but rather to make research visible and debatable, ensuring that experts are sufficiently informed to make sound judgments (Moed et al., 1985). Bibliometric and scientometric indicators are useful tools to help scientists and decision makers to obtain more objective information. They form an irreplaceable part in the scientific evaluation system, shedding light on the position of a part (country, region, institute, research group) in the whole, and this at any level.

ACKNOWLEDGEMENTS

Research for this article was supported by a grant from the project 'Investigation of the evaluation system of basic research' from the Chinese Ministry of Science & Technology. The authors thank M. Dekeyser (KHBO) for editorial help.

REFERENCES

Association of Chinese Publishers (2002). *China Publication Yearbook* (p. 840). Beijing: China Publication Yearbook Press (in Chinese).

Bonitz, M., Bruckner, E., Scharnhorst, A. (1999). The Matthew index — concentration patterns and Matthew core journals. *Scientometrics*, 44, 361–378.

China Science & Technology Monthly (1998). The comprehensive indicators of the China science and technology development. *China Science & Technology Monthly*, 3, 20–23 (in Chinese).

Chinese Basic Science (2003). Decision on improvement in the works of sci –tech evaluation. *Chinese Basic Science*, 3, 4–6.

Cyranoski, D. (2001). A great leap forward. *Nature*, 410, 10–12.

Démurger, S. (2001). Infrastructure development and economic growth: an explanation for regional disparities in China? *Journal of Comparative Economics*, **29**, 95–117.

Group for Statistics and Analysis of Chinese Articles (2001). Chinese sci –tech paper and citation database (CSTPCD): selection standards and application process. *Chinese Journal of Scientific and Technical Periodicals*, 12, 177–178 (in Chinese).

Jiang, Z. (2000). Science in China. *Science*, 288, 2317.

Jin, B., Li, L., Rousseau, R. (2003). *Production and productivity of Chinese scientists as a function of their age: the period 1995 –1999*. In G. Jiang, R. Rousseau, Y. Wu (Eds.), Proceedings of the 9th International Conference on Scientometrics and Informetrics (pp. 112–120). Dalian: Dalian Technological University Press.

Jin, B., Rousseau, R. (2001). An introduction to the barycentre method with an application to China's mean centre of publication. *Libri*, 51, 225–233.

Jin, B., Wang, B. (1999). Chinese Science Citation Database: its construction and applications. *Scientometrics*, 45, 325–332.

Jin, B., Wang, S., Wang, B., Rousseau, R., Wu, Z., Liu, X., Zhu, X. (1999). A unified method of counting international and domestic articles. *Journal of Management Sciences in China*, 2 (3), 59–65.(In Chinese with English abstract).

Jin, B., Zhang, J., Chen, D., Zhu, X. (2002). Development of the Chinese Scientometric Indicators (CSI). *Scientometrics*, 54, 145–154.

Li, J., Qian, G., Lam, K, Wang, D. (2000). Breaking into China. Strategic considerations for multinational corporations. *Long Range Planning*, 33, 673–687.

Liang, L. (2003). Evaluating China's research performance: how do SCI and Chinese indexes compare? *Interdisciplinary Science Reviews*, 28, 38–43.

Liang, L., Wu, Y., Li, J. (2001). Selection of databases, indicators and models for evaluating research performance of Chinese universities. *Research Evaluation*, 10, 105–113.

Meng, G., Wang, B. (1996). The library and information system of the Chinese Academy of Sciences. *Libri*, 46, 52–58.

Moed, H.F. (2002). Measuring China's research performance using the Science Citation Index. *Scientometrics*, 53, 281–296.

Moed, H.F., Burger, W.J.M., Frankfort, J.G., Van Raan, A.F.J. (1985). The use of bibliometric data for the measurement of university research performance. *Research Policy*, 14, 131–149.

National Bureau of Statistics & Ministry of Science and Technology (2002). *China Statistical Yearbook on Science and Technology* (p.5). Beijing: China Statistics Press.

Ren, S. (2002). Understanding and considerations of the SCI. *Chinese Journal of Scientific and Technical Periodicals*, 13, 5–8 (in Chinese).

Ren, S., Liang, P., Zu, G. (1999). The challenge for Chinese scientific journals. *Science*, 286, 1683.

Ren, S., Rousseau, R. (2002). International visibility of Chinese scientific journals. *Scientometrics*, 53, 389–405.

Russell, J., Rousseau, R. (2002). Bibliometrics and institutional evaluation. *Encyclopedia of Life Support Systems (EOLSS).* Part 19.3 Science and Technology Policy (Arvantis, ed.). Oxford (UK): Eolss Publishers.

Science & Technology Ministry of the P.R. China (2003). *The methods of S&T evaluation. Third edition*, Nov. 6, 2003 (Document no.2003 –308).

Shang, Y. (1988). China's scientific and technical standards as analysed from document statistics. *Journal of the China Society for Scientific and Technical Information*, 7, 252–266 (in Chinese). Translated in English and published in G. Rui, P. Dale (Eds.), *Information Science in China* (pp. 66–95). London: Aslib (1991).

Su, X. N., Han, X. M., Han, X. N. (2001). Developing the Chinese Social Science Citation index. *Online Information Review*, 25, 365–369.

Sun, Y., Xu, K.(2002). Neither underestimate the role of SCI, nor misuse it. *Chinese Journal of Scientific and Technical Periodicals*, 13,1–4 (in Chinese).

Tsui, K. (1996). Economic reform and interprovincial inequalities in China. *Journal of Development Economics*, 50, 353–368.

Wang, D. (2001). Some opinions on the role of SCI in *Research Evaluation*. *Chinese Journal of Scientific and Technical Periodicals*, 12, 292–294 (in Chinese).

Wang, Y., Wu, Y. (2001). *Status and trend of scientific and technical collaboration between People's Republic of China and Commonwealth of Australia: An analysis of scientific and technical papers co-authored by Chinese and Australians*. In F. Havemann, R. Wagner-Döbler, H. Kretschmer (Eds.), Proceedings of the Second Berlin Workshop on Scientometrics and Informetrics: Collaboration in Science and in Technology (pp. 211–218). Berlin: Gesellschaft für Wissenschaftsforschung.

Wu Y. (2002).Some facts about using SCI to measure research performance. *Chinese Journal of Scientific and Technical Periodicals*, 13, 39–41 (in Chinese).

Wu, Y., Liang L. (2001). Problems for attention and applying bibliographic quantitative indicators in evaluation of scientific research performance. *Chinese Journal of Scientific and Technical Periodicals*, 12, 110–112 (in Chinese).

Wu, Y., Pan, Y., Zhang, Y., Ma, Z., Pang, J., Guo, H., Xu, B., Yang, Z. (2003). *China Scientific and Technical Papers and Citations (CSTPC): history, impact and outlook*. In G. Jiang, R. Rousseau, Y. Wu (Eds.), Proceedings of the 9th International Conference on Scientometrics and Informetrics (pp. 352–361). Dalian: Dalian Technological University Press.

Yang, D.T. (1999). Urban–biased policies and rising income inequality in China. *American Economic Review*, 89, 306–310.

Chapter 23

DECOMPOSING NATIONAL TRENDS IN ACTIVITY AND IMPACT

A Study of Swedish Neuroscience Papers

Olle Persson and Rickard Danell

Inforsk, Department of Sociology, Umeå University, Umeå, Sweden
E-mail: Olle.Persson@soc.umu.se

Abstract: Publication and citation counts have become essential indicators for science policy in many countries. However, the overall national performance needs to be decomposed if it is to inform the development of appropriate and targeted policies. This can be accomplished by breaking down performance in terms of research institution, research groups and/or individual authors, as well as by applying different performance measures. In this chapter we show how the Swedish trend in activity and impact within neuroscience changes as we decompose trends according to actors and apply different measurements.

1. INTRODUCTION

National indicators of scientific publishing and impact are becoming increasingly important for informing national science policies. Most popular are trend series displaying activity and impact by sub-field, where sub-fields are based on a classification of journals into subject categories. Although these types of macro indicators give important signals about the international standing of a nation's research, they are only of limited use as a basis for national science policy making.

The policy implications of country level data are not obvious. For example, in a case of a country performing above world level would the appropriate policy response be to increase or to decrease spending? In principle, national performance above (or indeed below) world level could be used to motivate both an increase and a reduction in expenditure; we simply need more information to be able to design effective policies. We

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 515-528.

would, for example, want to know the extent to which the observed trends attributable to: differences between research institutions and scientists, changing publication practices, and type of data and measurement used.

Thus a number of more specific questions arise which call for a decomposition of national trends and a refinement of data and measurements. In this chapter we use an empirical study of papers in neuroscience journals focussing on Swedish publication activity and citation impact. Swedish neuroscience research shows a decline in relative citation impact during the last twenty years. In an earlier paper we tested the robustness of this trend by applying various types of measurements (Glänzel et al., 2003). We found that the negative trend did not change owing to journal coverage, changing publication behaviour, or if the citation impact was calculated relative to the journal impact or the sub-field citation rate. The study also indicated that the decline in relative impact could be located at the most productive authors and departments.

In this paper we will apply the decomposition approach to make a more detailed study of the role of specific actors. We will also study the effects of whole vs. fractional paper and citation counts, review papers, highly cited papers and highly cited authors. The purpose of this exercise is to reveal possible causes of the national trend, which is necessary for better informed and targeted science policies.

2. DATA AND METHOD

The data presented in this chapter are based on neuroscience papers from the 1986 to 2001 CD-ROM editions of the *Science Citation Index™ (SCI)*. The neuroscience set was defined by a list of 194 neuroscience journals, containing about half a million papers. Duplicate records and records lacking references or an address field were excluded. The study is based on the document types: articles, notes, letters, and reviews. All Swedish author addresses have been standardized.

Citation counts were generated from this set of neuroscience papers only, thus ignoring citations from journals outside this field, and by using a special search key. The search key consists of the last name of the first author, year, volume and starting page. The search generated about 3.3 million citations among the neuroscience papers.

Citation impact is defined relative to the world total and is called relative citation impact (RCI). RCI is calculated by dividing the mean citations per paper for Swedish neuroscience papers by the mean citation rate for all neuroscience papers.

Citations are summed over the whole period to make the RCI indicator comparable to the National Science Indicators (NSI) produced by the *Institute for Scientific Information* (ISI). The NSI data are widely referred to in Sweden, and the declining trend in RCI for Sweden which can be found in NSI data has recently become one of the main arguments for increased public spending on research.

3. WHOLE VERSUS FRACTIONAL COUNTING

There are several inflationary tendencies which need to be considered when analysing bibliometric data. Persson et al. (2004) have shown that authors, references, and citations grow faster than publications. The volume of papers produced and the number of citations received increase more or less automatically for any country if international collaboration and reference behaviour are not controlled. This makes the use of relative measures indispensable when studying trends in publication activity and citation impact.

We also need to study the effect of rapidly increasing collaboration amongst authors, institutions, and countries. In NSI data, as well as in other similar indicator sets, a whole paper is attributed to a country regardless of how many country names appear in the address field. For small countries, however, the share of internationally co-authored papers is high and increasing in almost all fields of science, thus about 40 percent of all SCI papers with a Swedish address are internationally co-authored with one or more countries. Since this share has increased by about one percent over the last 20 years, the use of whole counts will increasingly give a distorted picture of the actual publication output.

In Figure 23.1 we calculated trends in publication activity as a share of world total using whole and fractional counts. The two upper curves are based on whole counts, one from our study and the other from the National Science Indicators data. The trend is very similar and slightly negative for both of the whole count series.

When using fractional counts, however, the share of world total decreases and the trend becomes even more negative. This is obviously a reflection of the growing importance of international collaboration. The fractional count divides papers in proportion to the number of countries in the address field. If a paper is co-authored by Sweden and the US each country receives 0.5 papers. It would also be possible to fractionalise by the number of national addresses. For example, in the case of a paper with three US institutional addresses and one Swedish institutional address, Sweden would receive a count of 0.25. This latter type of fractional counting is likely to have effect of decreasing small nations′ share further. Whole counts may underestimate Sweden's decline of activity to the World.

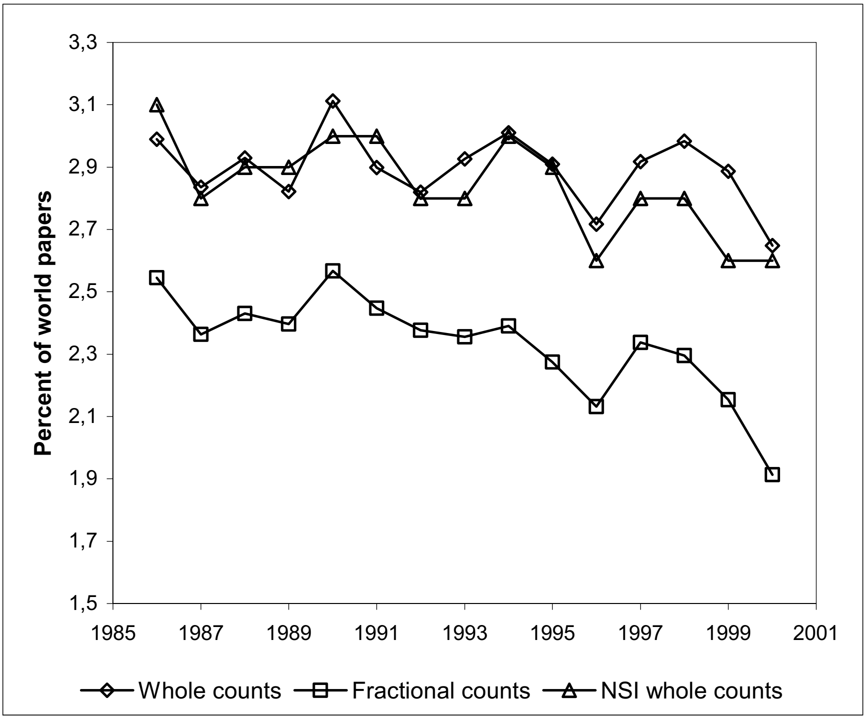


Figure 23.1. Whole vs fractional counts and publication activity of Sweden in neuroscience

The next question is whether using fractional counts also changes the Swedish trend in relative citation impact. To answer this we need a fractional relative citation impact measure. RCI for whole citation counts is based on the sum of citations, whilst the RCI for fractional citation counts is the sum of citation fractions, which are calculated by multiplying the paper fraction by the number of citations a paper receives.

Figure 23.2 displays the Swedish trend in terms of relative citation impact (RCI) using both whole and fractional citation counts. All trends are similar and negative, which means that fractionalisation, or taking account of international collaboration, does not change the Swedish trend or the level of citation impact. Had fractional RCI yielded a higher value than RCI based of whole counts, then one could surmise that Swedish domestic papers had a relatively higher citation rate than international papers with a Swedish address.

In order to study the effect of review articles, which usually have a higher citation impact, we also include a trend with review articles excluded. When this is done the peak for 1992 is gone, but the trend stays the same. However, the citation level is somewhat lower, indicating that Swedish review articles are more cited than articles in general. That review articles comprise about 3.5 percent of the Swedish output illustrates how sensitive the RCI level is to a few and highly cited papers.

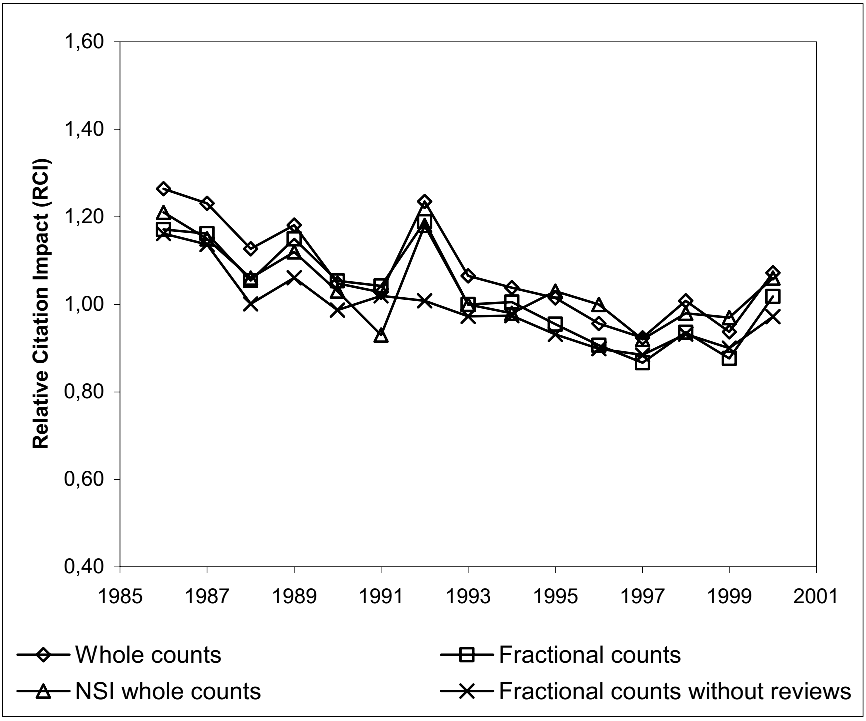


Figure 23.2. Relative citation impact based on whole versus fractional counting for Swedish papers in neuroscience

International collaboration has been found to have a strong positive effect on citation impact (Narin et al., 1990). This is supported by Figure 23.3, which shows that the RCI becomes much higher if domestic papers are eliminated. However, in 1990–1991 the level of RCI for domestic papers is not different from all Swedish papers.

It can be observed that collaboration between Sweden and the European Union (EU) has increased much faster than between Sweden and the US (*Web of Science*). Since US papers are on average more cited than European papers, Sweden's together with EU-countries could have the effect of reducing Sweden's relative citation impact. Figure 23.3, however, shows that removing neuroscience papers co-authored with the US, or with the EU, does not significantly affect the RCI level. To what extent this also holds for other sub-fields is not known.

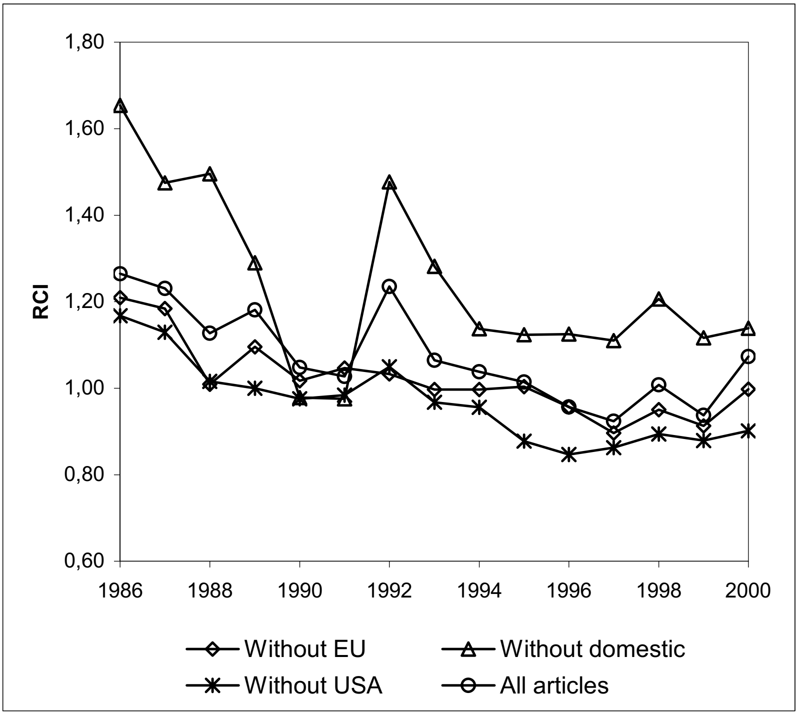


Figure 23.3. Type of international collaboration and relative citation impact for Swedish neuroscience papers (whole counts)

4. MAIN ACTORS IN SWEDISH NEUROSCIENCE

Sweden is a small country in terms of the world share of papers, yet there are about 200 different Swedish institutions producing about 9,000 neuroscience papers. However, the distribution of these papers by organisation is highly skewed. The five most productive institutions are all universities and account for 87 percent of the Swedish total. Karolinska Institute is by far the most productive organisation, followed by Lund and Gothenburg Universities. Karolinska's share is increasing strongly, whilst the other universities have a moderate growth (Figure 23.4).

Lund University and Karolinska have the highest relative citation impact, but all universities experience a downward trend in relative citation impact (Table 23.1). Since the decline in RCI is apparent for the most productive universities it can be concluded that they are accountable for the decline of the national impact trend.

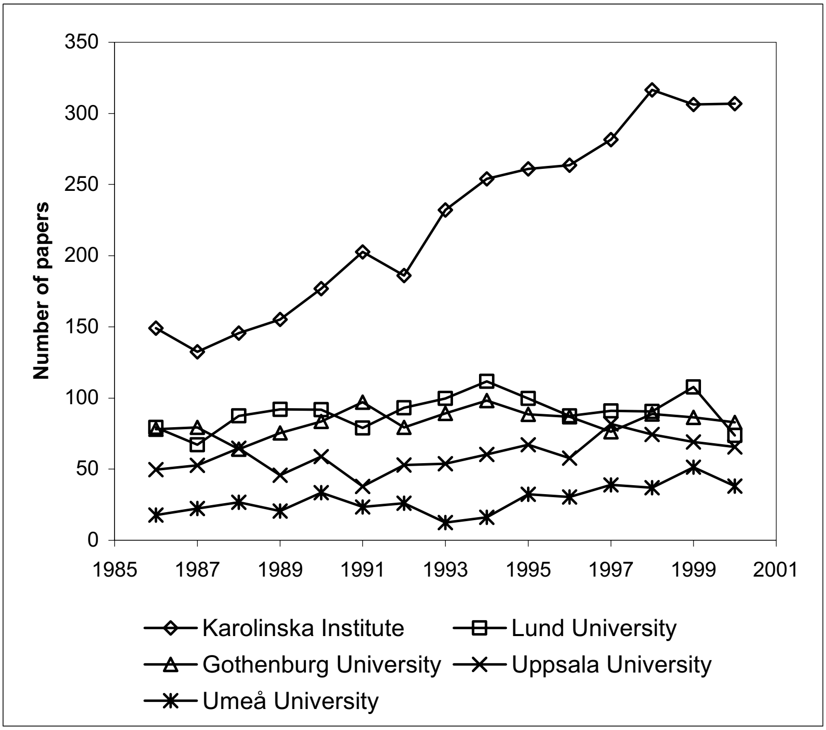


Figure 23.4. The main actors in Swedish neuroscience (fractional counts)

Table 23.1. Changes in relative citation impact for main Swedish organisations in neuroscience

Organisation	*Annual change of RCI 1986–2000*	*Average RCI for 1986–2000*
All Swedish papers	-0,02	1,08
Karolinska Institute	-0,02	1,26
Lund University	-0,04	1,32
Gothenburg University	-0,03	0,87
Uppsala University	-0,01	0,87
Umeå University	-0,01	0,96
Other organisations	-0,02	0,72

5. HIGHLY CITED PAPERS

Citation distributions are generally skewed to the advantage of a few highly cited documents. The same kind of distribution appears when citations are distributed by author, institution or country.

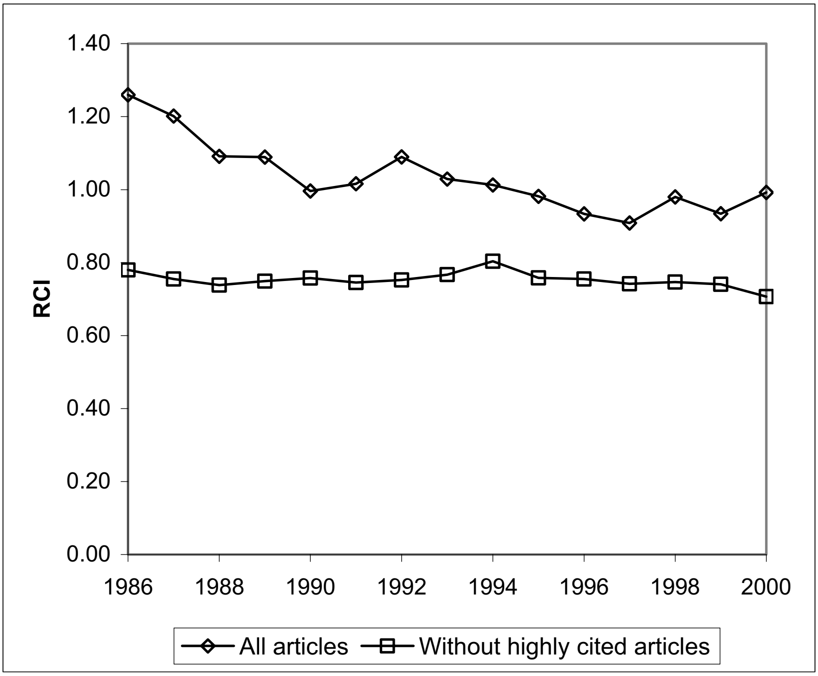


Figure 23.5. The effects of highly cited papers for the Swedish trend in neuroscience (whole counts, review articles excluded)

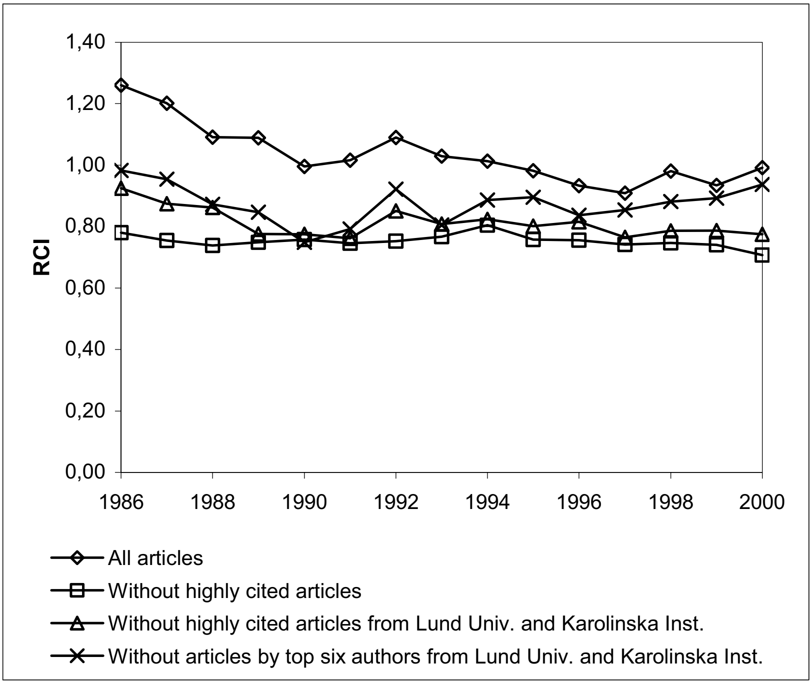


Figure 23.6. The effects of highly cited authors and universities for the Swedish trend in neuroscience (whole counts, review articles excluded)

Thus for Swedish neuroscience we could expect highly cited papers to play a major role for the national trend. In Figure 23.5 we have two curves. One with all papers included and one with the highly cited papers excluded.

To calculate the RCI for Sweden without the highly cited papers we excluded all Swedish papers belonging to the top five percent of the most cited papers in the World. Highly cited papers comprise 5.4 percent of the Swedish output over the period studied.

Since the citation trend, when the highly cited papers are excluded, is stable over time, the Swedish decline in relative citation impact can largely be attributed to a decline in highly cited papers. The important role of highly cited papers for citation trend analysis has been shown by Axnes and Sivertsen (2002) and Tijssen et al. (2002) showed that highly cited papers can be used for benchmarking scientific excellence.

In Figure 23.6 we study the effect of highly cited papers in more detail. The strongest change in the trend, and impact level, occurs when the highly cited papers from Lund University and Karolinska Institute are excluded. This is also what one would expect based on Table 23.1. Furthermore, if we exclude only twelve of the most productive authors of highly cited papers, six from Lund and six from Karolinska, then the level of impact sinks below the world total. This clearly demonstrates the influence of a few scientists on the impact of a whole country. The fact that the trend goes up in the last few years when these twelve individuals are excluded signals that there are new researchers or groups coming up.

Figure 23.7. Co-authorships amongst the highly cited Swedish neuroscientists

The size of the circles relates to the number of citations received. Grey circles represent authors who are among the top 20 for the years 1986–1995 as well as during 1996–1999. Black circles represent authors only present in the first period, and white circles those that appeared after 1995. The thickness of lines indicates the number of co-authorships.

6. HIGHLY CITED AUTHORS

It is quite obvious from what has been found so far that a few highly cited scientists account for the high Swedish impact at the start of the period. If we have a closer look at the most cited authors we can study the extent to which they collaborate and whether the network of collaboration is able to recruit new members and research groups.

To define highly cited authors we ranked authors by the sum of citations received during overlapping three year periods during 1986–1999. We excluded authors with less than 10 papers. For each three year period we made a top 20 list of the most cited authors. A total of 62 different authors appear in at least one of the top 20 lists. Next, we calculated the number of co-authorships among these top authors and produced a co-authorship map displayed in Figure 23.7.

We find that every one of the 62 most cited authors is connected to at least one other of the most cited authors via co-authorships. Most of the highly cited individuals are still active in the network (big grey circles), except for author 46 who is no longer active within the Swedish network. The capacity of the network to renew itself is indicated by as many as 15 authors (white circles) having made it to the top rank in the last four years.

7. PUBLICATION BEHAVIOUR

Since journals differ in their citation impact, journal impact is one factor which may affect citation trends. Figure 23.8 shows that the trend in RCI is not affected when we exclude the 25 most cited journals. The top neuroscience journals contain nearly 30 percent of all Swedish neuroscience papers. Thus we can conclude that the journals used for publication cannot explain the apparent decrease in Swedish citation impact. Glänzel et al. (2003) showed that Swedish papers are published in journals of equal impact to the world output.

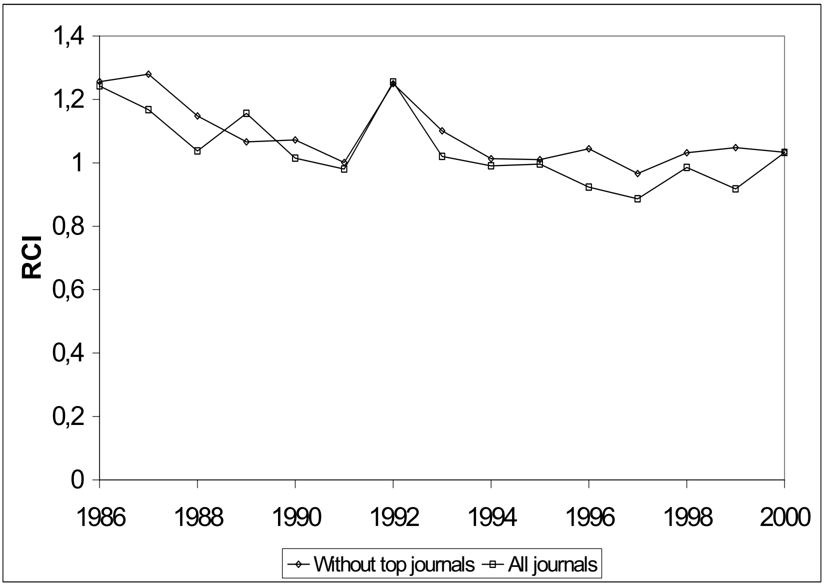


Figure 23.8. Journal impact and relative citation impact (whole counts, reviews included)

8. CONCLUSIONS

When decompose the Swedish trend in publication activity in neuroscience we have found that:

– Fractional counting reduces the share of Swedish papers.
– Fractional counting affects neither citation level compared to the World nor the decline in citation impact over time.
– Internationally co-authored papers are significantly more cited than domestic papers, but the level of citations or the negative trend is not sensitive to whether the collaborative partner is the EU or the US.
– The five most productive Swedish universities account for almost 90 percent of all Swedish neuroscience papers. There is a negative trend in relative citation impact for all of them.
– A decreasing share of the highly cited papers can explain the Swedish decline in relative citation impact.
– A relatively small group of authors accounts for the highly cited papers, and when their papers are excluded the Swedish citation impact is below world average and the decline is no longer visible.
– The most cited Swedish authors are all connected with each other by co-authorship links. A few authors with an exceptionally high number of

citations dominate the collaboration network, and they have been able to attract new collaborators with a high citation impact during the last few years.
- The trend in relative citation impact is not dependent on the impact of journals used for publication.

From the above list of observations we can conclude that a decomposition of the national trend yields additional information and which can feed into science policy making. The most striking result is that the high citation level of Swedish neuroscience is dependent on a very small group of scientists. On the other hand, the relative decline of Swedish impact can be attributed to a decreasing share of highly cited papers. These observations suggest that Sweden needs to produce research groups capable of creating research results of the same international standard as during the late 1980s. The constitution of these groups and the best way to identify, finance, and organize them are questions that bibliometrics alone cannot answer.

The added value of decomposing national indicators is that it answers questions which are frequently asked when scientists and policy makers are confronted with this type of macro indicator. One set of questions is about possible causes. For example, is a change in trend owed to better or declining research performance or to changing publication behaviour? For Swedish neuroscience we can conclude that decline is not an effect of changing publication behaviour, but rather a decline in research performance relative to other countries. Another set of questions is related to measurement. Especially for small countries, changing journal coverage and international collaboration may have strong effects on both publication counts and citation impact. Even though we could not find any major measurement effects in the case of Swedish neuroscience, this is often a source of uncertainty when interpreting national indicators.

ACKNOWLEDGEMENTS

The authors are very grateful to Dr Ingrid Schild at our own department for constructive suggestion comments.

REFERENCES

Aksnes, D.W., Sivertsen, G. (2001). *The effect of highly cited papers on national citation indicators*. 8th International Conference on Scientometrics and Informetrics, Vols 1 and 2 - Issi-2001, Proceedings**,** 23-30.

Glanzel, W., Danell, R., Persson, O. (2003). The decline of Swedish neuroscience: Decomposing a bibliometric national science indicator. *Scientometrics* 57 (2), 197-213.

Persson, O., Glanzel, W., Danell, R. (2004). Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics* , forthcoming.

Narin, F., Stevens, K., Whitlow, E.S. (1991). Scientific cooperation in Europe and the citation of multinationally authored papers. *Scientometrics* 21 (3), 313-323.

Tijssen, R. J. W., Visser, M. S., Van Leuwen, T.N. (2002). Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference? *Scientometrics* 54 (3), 381-397.

Chapter 24

NATIONAL PATTERNS OF TECHNOLOGY ACCUMULATION: USE OF PATENT STATISTICS

Lionel Nesta and Pari Patel

SPRU, The Freeman Centre, University of Sussex, UK
E-mail: L.J.J.Nesta@sussex.ac.uk

Abstract: We use US Patent Statistics to depict national patterns of technology accumulation in Japan and EU countries. Two properties of country profiles are confirmed, namely, *stability* over time with a country and *differentiation* across countries. The main novelty introduced here is the combined analysis of overall technological advantage, performance in fast growing areas and impact. The results show that in many areas of technology in which EU countries have an overall relative advantage, their performance in the sub-fields of highest technological opportunity is weak. On the other hand, Japan seems to have a consistent level of performance both in aggregate and in fast growing areas.

1. INTRODUCTION

The aim of this chapter is to review the considerable progress which has been made over the last 20 years in analysing national patterns of technology accumulation using patent statistics. Patent data provide a unique opportunity for social scientists interested in science and technology to understand and depict national and corporate patterns of technology accumulation. For example they have been used to make comparisons between countries and companies (US NSF Science and Engineering Indicators, European Science and Indicators Reports, Science & Technology Indicators published by Observatoire des Sciences et des Techniques in France) to analyse the nature and extent of globalisation of technology

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 531-551.

amongst firms (Cantwell, 1992; Patel and Pavitt, 1991) and to examine the linkages between science and technology (Narin and Olivastro, 1992).

In this chapter we provide a discussion of the use of patents statistics in describing certain aspects of national systems of innovations. In particular we analyse the similarities and differences amongst countries in national patterns of technology accumulation, in aggregate and in sectoral composition and specialisation. Thus we address the following issues: (i) the stability and similarities in sectoral patterns of technological profiles between technologically advanced countries; (ii) performance in fast growing areas of technology; (iii) the relationship between technological advantage and impact.

2. NATIONAL SYSTEMS OF INNOVATION

The notion of 'National Systems of Innovation' (NSI) is an useful one, since it treats explicitly what was ignored in earlier models of technical change: namely, deliberate 'intangible' investment in technological learning activities (Lundvall, 1992; Nelson, 1993). Such activities involve a variety of institutions, principally business firms, universities, other education and training institutions, and governments. An underlying rationale is that successful innovation, which is a necessary condition for improving welfare within a country, depends on the performance of these institutions as well as how well they interact as a system. The national systems approach is also a serious attempt at measuring and explaining the important differences between countries in the levels and patterns of investments in innovative activities.

Although a wide variety of studies are based on the NSI framework, there are inherent difficulties in defining and measuring the constituent parts of a national system of innovation. First, although the notion of NSI has gained tremendous momentum in science and technology policy research, there is a lack of common definitions across different approaches based on this notion (see Edquist, 1997). Second the unit of analysis is not always clear. Some studies take this to be the nation, others the region, and yet others an industrial sector. In the light of these difficulties the analysis of NSI is generally a descriptive and qualitative exercise. Most of the empirical work is concentrated on comparisons between different countries at a macro level, using data on R&D expenditures and patenting, combined with a description of the various institutions involved in national innovative activities (Lundvall, 1992; Nelson, 1993).

A limited number of studies have gone beyond this to measure and analyse the quantitative importance of different institutions and the interactions or knowledge flows between them (see Patel (1998) for a review). They have used a variety of S&T indicators (R&D, patenting, scientific publications, and strategic alliances) to examine: (i) inter-sectoral transactions embodying flows of technological knowledge (see the pioneering and unique work of Scherer, 1982); (ii) interactions among firms, primarily joint research activities and other technological collaborations; (iii) interactions among firms and institutions engaged in basic research such as universities and public research institutes.

Measurement of many of these interactions is the subject of other chapters in the Handbook. The purpose of this chapter is to explore the potential of patent statistics in mapping national patterns of technology accumulation. The underlying rationale is that such comparisons are a vital part of comparing different NSIs. They are a complement to many of the other qualitative and quantitative approaches depicting different aspects of the a system of innovation.

3. USE OF PATENT STATISTICS

A patent is first and foremost a legal instrument which gives a temporary monopoly to an inventor in exchange for detailed publication of the invention. Thus it allows the inventor to protect and profit from the invention and society to gain from wide dissemination of the knowledge about the invention. This information has long been used by academics and policy analysts to analyse trends in technology (see Schmookler, 1966). Over the last 20 years advances in information and communications technology have reduced the cost of storage and transmission of information and this has resulted in widespread use of patent statistics. Now all major patent offices provide online access to their data.

Publicly available patent documents contain a wealth of information which can be used in tracking advances in technology:

– Name and addresses of inventor(s);
– Name and addresses of the patent assignee(s);
– Date of filing;
– Technical class of the patent;
– References to other patents and scientific literature.

This information has been used in wide variety of studies to explore the nature, sources, and economic effects of technology. Apart from the subject of this chapter, i.e., comparisons between countries and between technical fields, the following issues have been addressed on the basis of patent data: the size distribution of innovating firms; their degree of internationalisation;

their technological diversification; the role played by technology in explaining international differences in export and productivity performance; and the links between technology and underlying basic research.

It has long been recognised that patents are an imperfect measure of the technological activities. However, the same is true of all technology indicators (R&D statistics and those based on innovation surveys). The main drawbacks of patent statistics are as follows:

First, there are major inter-sectoral differences in the relative importance of patenting in achieving its prime objective, namely, acting as a barrier to imitation. Thus recent studies show patenting to be relatively unimportant in automobiles but very important in pharmaceuticals (Arundel et al., 1995; Levin *et al.*, 1987; Bertin and Wyatt, 1988). Moreover, patents do not yet fully measure technological activities in software, since copyright law is often used instead as the main means of protection against imitation (see Samuelson, 1993). Given this inter-sectoral variety in the propensity to patent the results of R&D, patent statistics are most reliable when normalised by sectoral totals.

Second, there are major differences between countries in procedures and criteria for granting patents. For this reason comparisons are most reliable when using international patenting, or patenting in one country. US patenting statistics are a particularly rich source of information, given the rigour and fairness of criteria and procedures for granting patents, the strong incentives for firms to obtain patent protection for world class technology in the world's largest market (Bertin and Wyatt, 1988), and that these data are readily available. The same is increasingly true of the data from the *European Patent Office*. The illustrative analysis below is based on information from the USPTO, but the same analysis could be based on data from the EPO.

There is a further criticism of patenting as an indicator of technological activities which we think is not justified. We are not convinced that it is a drawback that patents differ greatly in their economic value (Schankerman and Pakes, 1986). The same is true of R&D projects (Freeman, 1982), and for the same reasons. Technological activities involve cumulative learning under uncertainty. There are therefore bound to be failures, major successes, and follow up improvements, all of which are interdependent. We would therefore expect similar and large variations in the distribution of the value of both R&D and patenting across all firms and countries.

4. SECTORAL PATTERNS OF NATIONAL TECHNOLOGICAL ACCUMULATION

In this section we discuss the use of patent statistics in making international comparisons of strengths and weaknesses in different technical fields amongst advanced industrialised countries. The underlying rationale for such comparisons is that they are a fundamental part of mapping any national system of innovation. They provide information on whether countries are specialising in leading edge technologies vital for introducing future products and processes or in those technologies that are likely to lead to few new opportunities in the future. Such information is important in assessing the future economic performance of a country.

4.1 Previous Studies

A pioneering study based on patent statistics for analysing the nature and determinants of technological accumulation amongst countries was that by Pavitt (1988). This seminal work is based on US patent data for 9 OECD countries across 29 areas of technology over the period 1963 to 1981. It includes one of the earliest discussions of the main determinants of the observed patterns of technological advantage. Perhaps the most comprehensive study covering most of the issues discussed below is Archibugi and Pianta (1992). The authors compare technological specialisation amongst US, Japan, and EU countries using both EPO and USPTO data. They use different sectoral classifications based either on International Patent Classification (IPC) or on Standard Industrial Classification (SIC). The results show that there are great differences across countries in their technological specialisation patterns. In general large countries are involved in a broad range of technical fields and smaller countries in a more narrow range.

Since the publication of these two pioneering studies a number of other scholars have used similar data and methodologies for examining a range of issues related to national technological accumulation. The purpose of the rest of this chapter is to present a more up to date illustration of the use of patent data in such an analysis.

4.2 Some Definitions

Patents can be assigned to a ‘country of origin’ on the basis of a number of different criteria: ‘priority’ country; address of the assignee; country of filing or address of the inventor. In the analysis below, and in common with most other studies, we have used the country address of the inventor[1].

The analysis of sectoral patterns of technological specialisations is based on a unique feature of patent statistics, namely, that they can be classified according to technology. In the comparisons below detailed classes of the US Patent Classification (USPC) have been aggregated to create 34 technological categories. There are a number of different classification schemes used in similar analyses. For example, Grupp (1992) presents an alternative based on IPC, consisting of 29 different technical fields.

One of the main problems in using US patent statistics is that they exaggerate the importance of US based inventors compared to inventors from other countries. The main reason being that the propensity to patent is always higher in the home country. For this reason the analysis below is based on foreign patenting in the US and all patents of US origin have been excluded.

For the analysis of fast growing sub-fields we have identified 1,500 (out of more than 70,000) detailed classes of the USPC with the highest absolute increase in foreign (non-US) patenting from 1971–1980 to 1991–2000. Their combined share increased steeply from 1.7% to 17.1% of all non-US patenting over the period. The underlying assumption is that fast growing fields reflect areas of greatest technological opportunity.

The analysis of impact is based on number of citations from the front page of each patent. The underlying assumption being that the higher the number of citations the greater the level of impact of a patent. Ideally such analysis should exclude ‘self-citations’ where the assignee (a company) name of the citing and the cited patent is the same. However, this requires the unification of all assignee names, which is a time (and resource) consuming process, not attempted here.

4.3 Indicators

Patterns of national technological accumulation can be compared on the basis of a number of different indicators; for example patent shares and growth rates of patenting. However the problem with many of these indicators is that they do not take into account the differing propensity to patent in the US amongst different countries. An indicator which corrects for this bias is the *Revealed Technological Advantage (RTA) Index*, which is defined as follows:

[1] Here we use information on the first inventor only, thus ignoring the problem of patents with multiple inventors in different countries.

$$RTA_{it} = \left(\frac{P_{it}}{\sum_t P_{it}} \right) \Bigg/ \left(\frac{\sum_i P_{it}}{\sum_{it} P_{it}} \right) \qquad (1)$$

This can be interpreted as an index of 'comparative advantage': with a value above unity indicating an area of relative strength and a value below unity an area of relative weakness. The definition of the index implies that its value is necessarily null or positive but is not bound by an upper limit[2]. For this reason some studies prefer to standardise the RTA measure. Sometimes such standardisation is performed by taking the logarithm of the index, which causes the RTA threshold value becomes zero. Positive values indicate areas of technological advantage and negative values indicate areas of technological disadvantage. There is no upper or lower limit to this standardised measure[3]. Alternatively one can force the RTA index to take values between -1 and +1 by computing the ratio of RTA minus one over RTA plus one: NRTA = (RTA - 1)/(RTA + 1). The threshold value remains zero, but the asymptotic limits are now[4] ±1. It is important to note that none of these transformations will affect the rank of a series of RTA values. In the following analysis, for the sake of simplicity, we use the simple RTA measures defined in Eq. (1).

The analysis of fast growing fields is based on the *Fast growing Specialization Index (FGSI)*, which is simply the share of a country in the fast growing sub-fields of a technology divided by its share of all patents in that technology:

$$FGSI_{it} = \left(\frac{F_{it}}{\sum_t F_{it}} \right) \Bigg/ \left(\frac{P_{it}}{\sum_t P_{it}} \right) \qquad (2)$$

where F_{it} indicates the number of patents held by country i in fast growing technology areas belonging to technology class t. The *FGSI* ratio can be interpreted as an index of 'comparative advantage in rapidly changing technologies': with a value above unity indicating an area of relative

[2] Where RTA $\in$ [0 ; +∞].
[3] I.e. log(RTA) $\in$ [-∞ ; +∞].
[4] I.e. NRTA $\in$ [-1 ; +1].

strength and a value below unity an area of relative weakness. Our assumption is that technologies grow rapidly in the early stage of their development so that countries with FGSI above unity are likely to benefit substantially from a 'first mover' advantage over other countries.

The third indicator analysed below is the *Relative Impact Index (RII)*, defined as the citations per patent for a particular country in a technology divided by aggregate citations per patent in that technology[5]:

$$RII_{it} = \left(\frac{C_{it}}{\sum_t C_{it}} \right) \Bigg/ \left(\frac{P_{it}}{\sum_t P_{it}} \right) \quad (3)$$

where C_{it} indicates the number of citations to patent received by country *i* in technology class *t*. Again, a value above (below) unity signifies that the country has relatively high (low) impact of patents.

An interesting feature of these three measures is that they can be used to calculate further indices that characterise patterns of specialisation within or across countries. To determine whether a country has established niches of technological excellence or broadened its national technological competences to a wider spectrum, one can calculate the coefficient of variations of any of the above measures. In the case of RTA indices this yields:

$$CV_i = \frac{\sigma_{\mathrm{RTA}_{it}}}{\mu_{\mathrm{RTA}_{it}}} \quad (4)$$

where, for a given country i, the CV_i is the coefficient of variation of the RTA, σ and μ are respectively the standard deviation and arithmetic mean of RTA values. This is a measure of the concentration of patent counts across technologies: a high CV means that the country is concentrating its areas of excellence within a narrow band of technological competences. Conversely a low CV means that the country is developing its competences uniformly across the range of technologies. Thus, Eq. (4) provides information on the degree of technological specialisation within a country. Eq. (4) can easily be extended to CV_t which measures, per technology, the concentration of patent counts across countries. This tells us whether a given technology class is concentrated within one or a few countries or whether it is more diffused. In

[5] Again, US citations have been excluded from the analysis.

other words this measures the degree of diffusion of a given technology across countries, and thus helps us characterise whether a technology is common across a broad range of countries (low CV_t) or whether it constitutes a niche technology (high CV_t).

4.4 Sectoral Patterns of Technological Advantage

Table 24.1. RTA in Selected Advanced OECD Countries: 1991–2000

Technology	*DE*	*FR*	*UK*	*IT*	*JP*	*NL*	*SE*	$CV_t \times 100$
Inorganic Chemicals	1.39	2.04	0.92	1.05	0.70	1.45	0.96	74.3
Organic Chemicals	1.83	1.10	1.05	1.63	0.85	1.38	0.27	56.0
Agricultural Chemicals	2.52	0.68	1.62	0.46	0.63	0.54	0.12	103.4
Chemical Processes	1.25	1.25	1.43	1.06	0.86	1.28	0.85	38.4
Hydroc. Min. Oils, etc.	1.04	3.32	1.85	2.15	0.44	2.14	0.27	89.2
Bleaching & Dyeing	2.10	2.63	1.32	1.37	0.37	0.46	0.84	132.7
Drugs & Bioengineering	1.06	1.93	2.28	1.65	0.57	1.24	1.29	59.2
Plastic & Rubber Product	1.14	1.02	0.89	1.20	1.06	1.01	0.72	177.5
Materials	0.84	0.89	0.82	0.53	1.34	0.68	0.60	56.8
Food and Tobacco	0.82	1.17	1.92	1.69	0.58	4.16	1.29	66.8
Metal Treatment	0.85	1.06	0.81	0.48	1.12	0.59	1.61	65.0
Apparatus for Chemicals	1.33	0.98	0.84	1.63	0.71	1.08	1.55	51.4
Gen. Non-Elec Ind Eq.	1.63	1.23	1.07	1.04	0.74	0.56	1.42	58.7
Gen. Elec Ind Apparatus	0.90	0.89	0.71	0.83	1.15	0.63	0.68	44.4
Non-Elec Spec Ind Eq.	1.68	1.10	0.91	2.16	0.56	0.93	1.61	36.4
Metal Working Eq.	1.43	0.80	0.79	1.67	0.87	0.82	1.63	99.0
Ass. & Material handling app.	1.52	0.61	0.67	2.57	0.79	1.05	1.65	56.5
Induced Nuclear Reactions	1.32	4.49	0.64	0.58	0.72	0.60	4.68	180.1
Power Plants	1.37	2.24	1.83	0.42	0.71	0.66	1.59	78.9
Road Vehicles and Engines	1.53	0.37	0.59	0.65	1.27	0.18	0.96	89.4
Other Transport Eq.	1.42	1.35	0.67	1.27	0.66	0.50	1.31	77.6
Aircraft	1.71	4.99	2.84	0.65	0.15	0.78	0.80	122.7
Mining & Wells Mach & Proc	0.80	2.16	4.36	0.40	0.12	1.43	2.02	200.3
Telecommunications	0.69	1.32	1.15	0.49	1.01	1.26	2.38	82.9
Semiconductors	0.29	0.40	0.24	0.68	1.32	0.71	0.30	126.5
Elec. Devices and Systems	0.89	1.00	0.66	0.57	1.02	1.56	0.75	59.6
Calculators, Computers, etc.	0.42	0.70	0.79	0.71	1.42	0.67	0.51	64.8
Image and Sound Eq.	0.23	0.35	0.52	0.14	1.63	1.10	0.13	100.8
Photography and Photocopy	0.36	0.10	0.49	0.26	1.72	0.25	0.04	188.9
Instruments and Controls	1.02	0.99	1.15	0.59	1.08	1.05	0.79	30.5
Miscellaneous Metal Products	1.36	1.16	1.16	1.28	0.48	0.81	1.25	43.2
Textile & Wood Products	0.96	1.64	1.04	3.03	0.41	0.48	1.36	57.8
Dentistry and Surgery	1.19	1.34	1.43	1.39	0.55	1.74	3.29	54.8
Other	0.81	1.14	1.31	0.98	0.43	1.39	1.43	48.1
Number of RTAs above unity	*21*	*22*	*17*	*17*	*12*	*16*	*17*	
Specialization ($CV_i \times 100$)	*43.2*	*75.5*	*65.3*	*62.7*	*47.4*	*68.3*	*76.7*	

Table 24.1 compares 7 advanced OECD countries in 34 technological fields on the basis of the RTA measure (defined above) in the period 1991–2000. Table 24.1 also displays information on the number of RTAs above unity, the degree of specialisation of country i (CV_i) and the degree of technological commonality across countries (CV_t).

Looking first at the broad picture we note that Germany and France have leading positions in terms of the number of RTAs above unity (21 and 22 respectively), signifying the breadth of their national knowledge base. All other European countries have a significant number of technologies in which they have developed a critical mass of technological competences (16 or more with RTA > 1). The degree of specialisation varies greatly by country with Germany and Japan having the lowest values compared to the other 5 countries. On the basis of a much larger sample of countries Archibugi and Pianta (1992) show that this is partly a function of the size of the country. The position of Japan is partially explained by it patenting significantly more in the US than any European countries. In this case the RTA measure converges asymptotically towards unity.

The last column of Table 24.1 provides the coefficient of variation of RTA values per technology (CV_t) and reveals some interesting patterns. One of the highest CV_t is in *Photography and Photocopy* (CV = 188), where Japan has a clear technological advantage (RTA = 1.72), and where the European countries are weak (no RTA value exceeds 0.5). Amongst the other niche technologies are *Mining and Machinery* and *Nuclear*. In the case of the former the explanation is that only those countries with abundant raw materials would be involved in developing competences in this area. In the case of the latter only a few countries have decided to take the nuclear route for the provision of electricity. France and the UK have a strong technological advantage in *Aircraft* (RTA = 4.99 and 2.84 respectively, and CV = 123), confirming their leading involvement in the Airbus industry and their leading role in the development of aircraft technology in general.

The preceding analysis implies that some countries face major barriers in building a significant national technology base in certain areas. For example, the position of France in *Photography and Photocopy* (RTA = 0.10) and in *Image and Sound Equipment* (RTA = 0.35) makes it more difficult for France to play a significant role in these areas in the future. Likewise the poor performance of Japan in *Aircraft* and *Pharmaceuticals* related technologies shows that it is difficult to build high levels of national competence in a wide range of high technology areas simultaneously.

The evolution of sectoral patterns of technological advantage can be analysed by comparing RTAs over time. Tables 24.2a and 24.2b give examples of such analysis for Germany and the UK over the period 1971–

2000. There are some elements common to both countries. For example, they are both relatively strong and becoming stronger in most of the Chemical related technical fields and in Aircraft, and weak and getting weaker in fields related to IT (*Semiconductors*, *Computers*, *Radio and TV*, *Photography and Photocopy*). In terms of differences Germany has increasing relative strengths in many of the machinery related technologies and motor vehicles, which are areas of weakness for the UK. On the other hand the UK has relative advantage in two ‘high-tech’ areas where Germany has a disadvantage: *Drugs and Telecommunications*.

Table 24.2a. Evolution of Technological Advantage 1971—2000: Germany

	Increasing	**Stable**	**Decreasing**
Advantage	Inorganic Chemicals Organic Chemicals Agricultural Chemicals Chemical Processes Bleaching & Dyeing Plastic & Rubber Products Apparatus for Chemicals, etc. General Non-Elec. Eq. Specialised Non-Elec. Eq. Metallurg. Working Eq. Ass. Material Handling App. Power Plants Road Vehicles & Engines Other Transport Eq. Aircraft Misc. Metal Products	Induced Nuclear Reactions	
Disadvantage	Hydrocarb., Mineral Oils, etc. Drugs & Bioengineering Food & Tobacco	Materials Metallurg. & Metal Treatment General Elec. Industrial App. Mining, Wells Mach. & Proc. Electrical Devices & Systems Instruments & Controls	Telecommunications Semiconductors Computers & Office Eq. Image & Sound Equipment Photography & Photocopy

Table 24.2b. Evolution of Technological Advantage 1971–2000: UK

	Increasing	Stable	Decreasing
Advantage	Organic Chemicals Agricultural Chemicals Chemical Processes Hydrocarb., Mineral Oils, etc. Bleaching & Dyeing Drugs & Bioengineering Food & Tobacco Aircraft Mining, Wells Mach. & Proc. Instruments & Controls	Power Plants Telecommunications Misc. Metal Products	Plastic & Rubber Prod. General Non-Elec. Eq. Induced Nuclear Reactions
Disadvantage	Inorganic Chemicals Photography & Photocopy	Metallurg. & Metal Treatment Apparatus for Chemicals, etc. Specialised Non-Elec. Eq. Textile & Wood Products	Materials General Elec. Industrial App. Metallurg. Working Eq. Ass. Material Handling App. Road Vehicles & Engines Other Transport Eq. Semiconductors Electrical Devices & Systems Computers & Office Eq. Image & Sound Equipment

The RTA index is *increasing* or *decreasing* when it changes by more than 10% in the period 1971–80 to 1991–2000, otherwise it is *stable*. A country has an advantage in a technology when its RTA is greater than unity for the whole period 1971–2000.

4.5 Stability and Similarities amongst Countries

Table 24.3 examines the similarities and differences between countries' technological specialisations in greater and more systematic detail. It uses correlation analysis to measure both the stability over time of each country's sectoral strengths and weaknesses in technology (first row), and the degree to which they are similar to those of other countries (correlation matrix). The first row shows that all 7 countries have a statistically significant degree of stability in their technological specialisations between the 1970s and the 1990s, confirming the path–dependent nature of national patterns of accumulation of technological knowledge. There are some differences across countries in the degree of stability with Germany, France, and Japan having a much more stable profile than Netherlands, UK and Sweden.

The correlation matrix confirms the presence of differentiated technological profiles. For example, it shows that Japan is negatively correlated with all European countries. This means that Japan has accumulated competences in technological areas which are different from those of European countries. However, the six European countries do not form a homogeneous block: only 3 coefficients are positively and significantly correlated out of 15. The correlated profiles are those of France and the UK (correlation = 0.494), France and Sweden (correlation = 0.398) and the UK and the Netherlands (correlation = 0.333). Altogether, these three correlations represent only 14% of the 21 correlations between pairs of countries in Table 24.3. This low level of correlations suggests that countries differ greatly in their profiles of technological competences.

Table 24.3. Stability and Similarities Amongst Countries in their Sectoral Specialisations: Correlations of RTA Indices across 34 Sectors

Stability: Correlations Over Time: 1971—80 to 1991—2000							
	DE	*FR*	*UK*	*IT*	*JP*	*NL*	*SE*
	0.80*	0.87*	0.61*	0.74*	0.89*	0.51*	0.63*
Similarities: Correlations Amongst Countries: 1991–2000							
	DE	*FR*	*UK*	*IT*	*JP*	*NL*	*SE*
FR	0.29						
UK	0.14	0.49*					
IT	0.28	0.05	-0.03				
JP	-0.54*	-0.62*	-0.67*	-0.46*			
NL	-0.18	0.07	0.33*	0.25	-0.28		
SE	0.08	0.40*	0.11	0.09	-0.36*	0.06	

* Denotes correlation coefficient significantly different from zero at the 5% level.

4.6 Assessing Country Performance in Fast Growing Technologies

The analysis of fast growing areas provides us with an indication of how well a country is performing in technologies with the highest level of technological opportunities for the future. A large proportion of the fast growing areas can be found in IT related technologies and in Pharmaceuticals and Biotechnology. A comparison of *FGSI* (defined above) provides information on the position of countries in technologies that will in the future become very important components of the production system for firms and industries[6]. A key question in this analysis is the extent to which overall specialisation in a technology is related to performance in the fastest growing areas of the technology. One way of assessing this is by plotting

[6] Annex 1 presents the results for the seven OECD countries in a similar fashion to that in Table 24.1.

each country on a 2-dimensional map with *FGSI* along the X-axis and *RTA* along the Y-axis as shown in Figure 24.1.

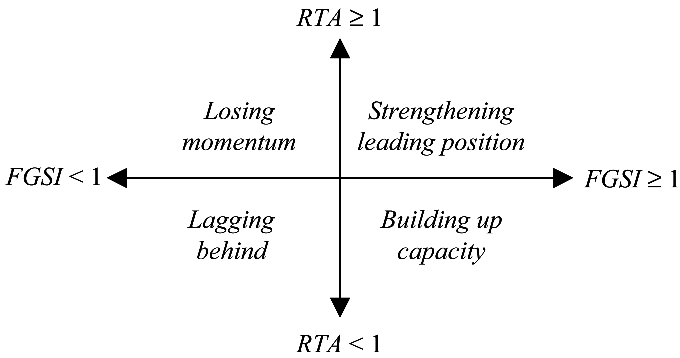


Figure 24.1. Technology map of countries

Countries located in the upper right quadrant for a particular technology have a strong advantage and they exhibit a high level of specialisation in fast growing areas, putting them in a position to reinforce their advantage over time. Those in the lower right quadrant have a low overall technological advantage but exhibit specialisations in fast growing areas: such countries are building up national technological competences in key technologies of the future. Countries belonging in the upper left quadrant in a specific technology have a high overall advantage but have a low specialisation in fast growing areas: it is likely that they will lose momentum in the future. Finally those located in the lower left quadrant have a low technological advantage and a low specialisation in fast growing areas: such countries are simply lagging behind.

Figures 24.2 to 24.4 present plots for each of the major technological families: Chemical, Electrical and Mechanical. Figure 24.2 shows the overall strength of EU countries in chemical related technologies. Netherlands is in a leading position in *Inorganic Chemicals* and *Food and Tobacco*. France also has a leading position in 3 areas of chemical related technologies, including *Drugs and Bioengineering*. There are some signs that the two leading EU countries, Germany and the UK, are losing momentum in a range of different chemical technologies, including *Drugs and Bioengineering*. Thus although they have overall advantage in many technical fields, their performance in fast growing fields is weak. At the same time Figure 24.2 shows that Japan is building up capacity in chemical technologies.

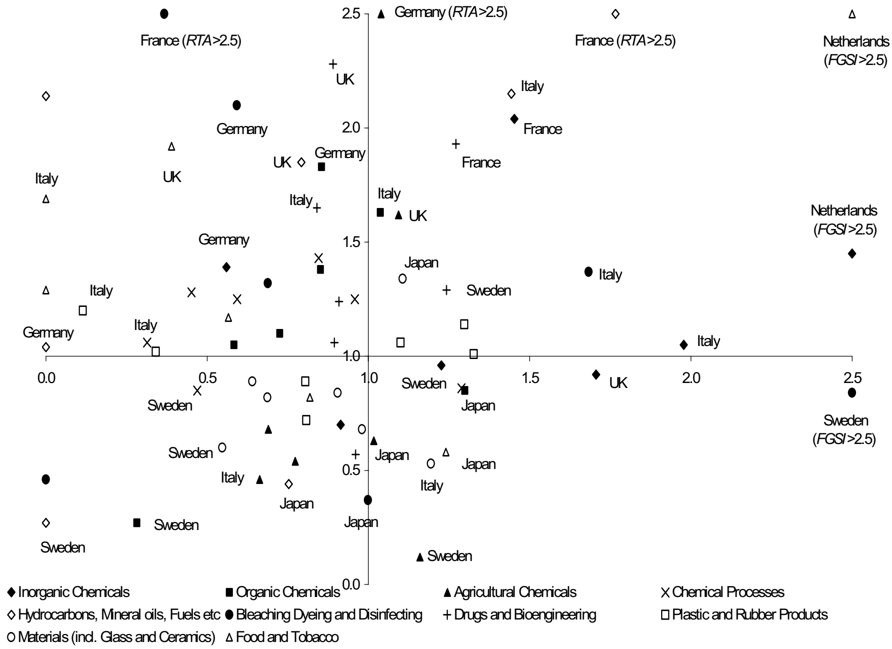


Figure 24.2. Technology map of countries: Chemical related technologies (1991–2000)

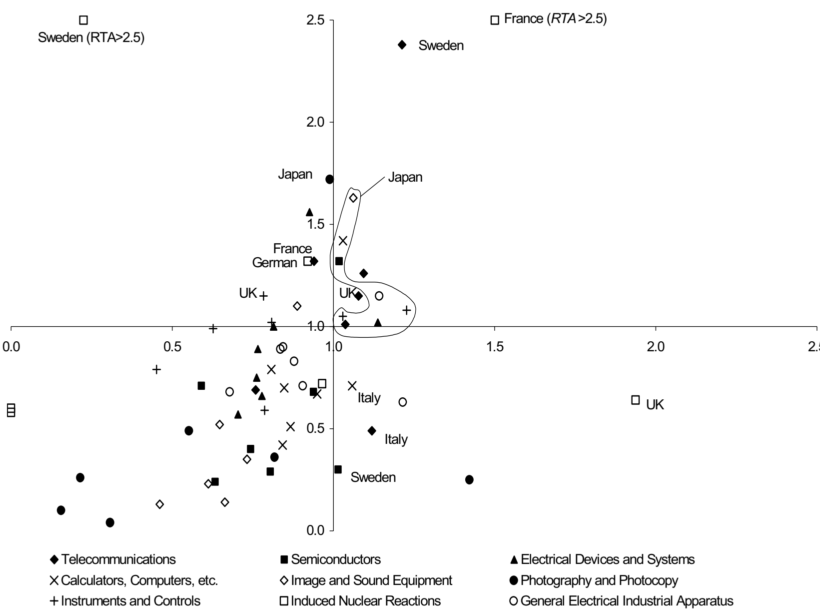


Figure 24.3. Technology map of countries: Electrical related technologies (1991–2000)

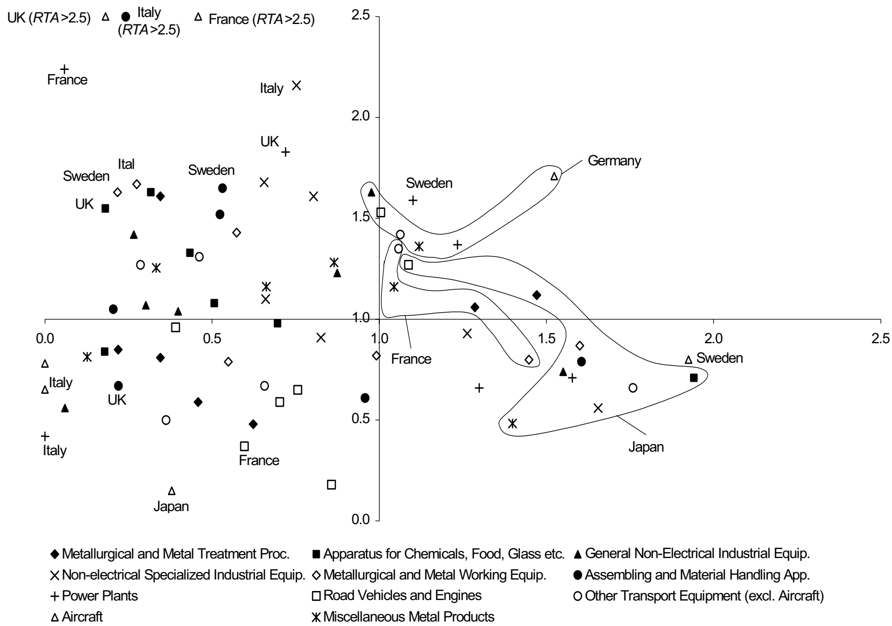


Figure 24.4. Technology map of countries: Mechanical related technologies (1991–2000)

Figure 24.3 shows that Japan is in a leading position in many of the electrical related technologies. In particular it has both a high technological advantage and a high level of performance in fast growing areas in many of the IT related areas: *Computers*, *Semiconductors*, and *Image and Sound*. In general EU countries perform badly in most areas of electrical technologies. They are losing momentum in a broad range of technical fields. A major exception is *Telecommunications* technologies, where Sweden, Netherlands, and the UK are in a leading position, and Italy is building up capacity. France also has a strong advantage in *Nuclear* technologies.

Finally Figure 24.4 presents the map for Mechanical related technologies. It shows that Germany is in a very strong position in a wide range of technical fields: it has both a high overall advantage and good performance in fast growing areas. Two areas where this is especially the case are *Aircraft* and *Power Plants*. France is also in a leading position in a number of mechanical technologies. One of the surprising results is that although the two leading EU countries in *Aircraft* technologies, the UK and France, have a high overall advantage, their performance in fast growing areas is weak. The other surprise is that Italy has no technologies in either the *leading* or *building capacity* categories. Figure 24.4 also shows that Japan is *building* *capacity* in a range of mechanical technologies, especially *Chemical Apparatus*, *Other Transport* and *Specialised Machinery*. A problem for EU countries is that in automobile related technologies they are lagging behind and this is an area in which Japan is in a lead position.

4.7 Relationship between Country Profiles and Impact

The final issue addressed in this chapter is the relationship between the profiles of countries according to the 4-way classification employed in Figures 24.1 to 24.4 and a measure of impact, namely the Relative Impact Index (RII) as defined above. In other words we address the question whether there are statistically significant differences in terms of impact when each technology—country combination is aggregated according to the 4 quadrants in Figure 24.1. The results based on an analysis of variance of RII (reported in Table 24.4) show that in technologies where countries are either strengthening their leading position or building up capacity, they have a statistically significant higher level of impact than those technologies in which they are losing momentum or lagging behind. This implies that there is a stronger relationship between performance in fast growing fields and impact than between overall technological advantage and impact.

Table 24.4. Relationship between Technology Profiles and Impact: Analysis of variance

	N.Obs.	*Average RII*	*Std. Dev. Of RII*	*95% Confidence Interval for Mean*	
Strengthening leading position	42	0.999	0.140	0.955	1.042
Building up capacity	32	1.011	0.288	0.907	1.115
Losing momentum	75	0.921	0.149	0.887	0.955
Lagging behind	82	0.846	0.247	0.792	0.901
Total	231	0.921	0.218	0.893	0.949

F-stat=7.346. Critical Probability Value $P(F > F^{0.01}_{3,200}) = 0.00$. Significant at 1% level.

5. CONCLUSIONS

This chapter has provided the methodological basis for the potential application of patent data at the macro level, giving examples of possible uses of patent statistics in analysing national patterns of technology accumulation. The analysis presented above confirmed that such patterns are persistent over time, implying that path dependency is a key feature of technology accumulation. It also showed that there is a great deal of variety in technology profiles across countries. Japanese technological competences are very different from those of European countries. However, Europe as a whole is far from homogeneous.

A key innovation of the approach outlined in this chapter is the combined analysis of overall technological advantage, performance in fast growing areas and citation performance. The results show that in many areas of technology in which EU countries have an overall relative advantage, their performance in the sub-fields of highest technological opportunity is weak. On the other hand, Japan seems to have a consistent level of performance both in aggregate and in fast growing fields. Finally, the analysis also shows that there is a stronger relationship between performance in fast growing fields and impact than between overall technological advantage and impact.

It is clear from the above discussion that patent statistics present a unique opportunity to understand and depict national patterns of technology accumulation. However such analysis is only the starting point in analysing national systems of innovation. A key issue is how to explain the major inter-country differences in the evolution of patterns of technological advantage. In earlier work (Patel and Pavitt, 1994) we have argued that a country's sectoral technological advantage can come from two sources. First, there are country-specific inducement mechanisms to which local firms have preferential access: chief amongst these will be access to defence and other public sector markets, to abundant raw materials, and to idiosyncratic consumer tastes. Second, there are firm- or region-specific skills that enable firms to respond to a variety of technological opportunities and market needs: chief amongst these will be their own in-house mastery of one or more of the three pervasive ‘technological families’: electrical–electronic, chemical, and mechanical (and related automobile technology). Using such a framework to interpret some of the results outlined above would lead to a much richer analysis of national systems of innovation.

Other promising possibilities involve relating patent data with other types of systematic information related to key features of a national system of innovation such as national R&D budgets and availability of venture capital. Another enhancement could be to relate the patterns identified above with descriptive information such as public programmes in technology policy and public–private relationships.

REFERENCES

Archibugi, D., Pianta M. (1992). *The technological specialization of advanced countries*, Kluwer, Dordrecht.

Arundel, A., van de Paal, G., Soete, L. (1995). *Innovation strategies of europe's largest industrial firms,* PACE Report, MERIT, University of Limbourg, Maastricht.

Bertin, G., Wyatt, S. (1988). *Multinationals and industrial property*, Harvester-Wheatsheaf, Hemel Hempstead.

Cantwell, J. (1992). *The internationalisation of technological activity and its implications for competitiveness*. In Granstand, O., Hakanson, L., Sjolander, S. (Eds.), Technology Management and International Business: Internationalisation of R&D and Technology. Chichester: Wiley.

Edquist, C. (1997). *Systems of innovation: technologies, institutions and organizations*. London: Pinter Publishers/Cassell Academic.

Freeman, C. (1982). *The economics of industrial innovation*. London: Pinter.

Grupp H. (1992). *Dynamics of science based innovation*. Heidelberg: Springer-Verlag.

Levin, R., Klevorick, A., Nelson R., Winter S. (1987). Appropriating the returns from industrial Research and Development. *Brookings Papers on Economic Activity*, 3, Washington.

Lundvall, B-Å. (1992). *National systems of innovation: towards a theory of innovation and interactive learning*. London: Pinter.

Narin, F., Olivastro D. (1992). Status report: linkage between technology and science, *Research Policy*, 21, 237–249.

Nelson, R.R. (Ed.). (1993). *National systems of innovation: A comparative study*, Oxford: Oxford University Press.

Patel, P. (1998). Indicators for systems of innovation and system interactions: technological collaboration and inter-active learning, *IDEA Paper Series*, 11. Oslo: STEP Group, 20p.

Pavitt, K. (1988). *International patterns of technological accumulation*. In N. Hood, J.-E. Vahlne (Eds.), Strategies in Global Competition. London: Croom Helm.

Samuelson, P. (1993). A case study on computer programs. In J. Wallerstein, M. Mogee, R. Schoen (Eds.), *Global dimensions of intellectual property rights in Science and Technology*. Washington, DC: National Academy Press.

Schankerman, M., Pakes, A. (1986). Estimates of the value of patent rights in european countries during the Post-1950 Period, *Economic Journal*, 96, 1052–1076.

Scherer, F. M. (1982). Inter-industry technology flows in the United States, *Research Policy*, 11, 227–245.

Schmookler, J. (1966). *Invention and Economic Growth*. Cambridge (Mass.): Harvard UP.

APPENDIX

Table 24.A1. FGSI in Selected Advanced OECD Countries: 1991–2000

Technology	*DE*	*FR*	*UK*	*IT*	*JP*	*NL*	*SE*	*CV_t×100*
Inorganic Chemicals	0.56	1.45	1.71	1.98	0.91	3.16	1.23	129.0
Organic Chemicals	0.85	0.73	0.58	1.04	1.30	0.85	0.28	94.5
Agricultural Chemicals	1.04	0.69	1.09	0.66	1.02	0.77	1.16	98.4
Chemical Processes	0.96	0.59	0.85	0.31	1.29	0.45	0.47	77.0
Hydroc. Min. Oils, etc.	0.00	1.77	0.79	1.44	0.75	0.00	0.00	315.7
Bleaching & Dyeing	0.59	0.37	0.69	1.68	1.00	0.00	5.82	187.4
Drugs & Bioengineering	0.89	1.27	0.89	0.84	0.96	0.91	1.24	29.4
Plastic & Rubber Product	1.30	0.34	0.80	0.11	1.10	1.33	0.81	150.0
Materials	0.90	0.64	0.69	1.19	1.11	0.98	0.55	97.7
Food and Tobacco	0.82	0.57	0.39	0.00	1.24	3.15	0.00	173.7
Metal Treatment	0.22	1.29	0.35	0.62	1.47	0.46	0.35	120.9
Apparatus for Chemicals	0.43	0.70	0.18	0.32	1.94	0.51	0.18	119.8
Gen. Non-Elec Ind Eq.	0.98	0.87	0.30	0.40	1.55	0.06	0.27	107.2
Gen. Elec Ind Apparatus	0.84	0.84	0.91	0.88	1.14	1.22	0.68	52.2
Non-Elec Spec Ind Eq.	0.66	0.66	0.83	0.75	1.65	1.26	0.80	103.5
Metal Working Eq.	0.57	1.45	0.55	0.27	1.60	0.99	0.22	109.1
Ass. & Material handling app.	0.52	0.96	0.22	0.24	1.61	0.20	0.53	108.5
Induced Nuclear Reactions	0.92	1.50	1.94	0.00	0.97	0.00	0.22	180.7
Power Plants	1.23	0.06	0.72	0.00	1.58	1.30	1.10	127.1
Road Vehicles and Engines	1.00	0.60	0.70	0.76	1.09	0.86	0.39	77.0
Other Transport Eq.	1.06	1.06	0.66	0.29	1.76	0.36	0.46	100.7
Aircraft	1.52	0.46	0.18	0.00	0.38	0.00	1.93	283.4
Telecommunications	0.76	0.94	1.08	1.12	1.04	1.09	1.21	42.2
Semiconductors	0.80	0.74	0.63	0.94	1.02	0.59	1.01	62.4
Elec. Devices and Systems	0.77	0.81	0.78	0.70	1.14	0.93	0.76	49.4
Calculators, Computers, etc.	0.84	0.85	0.81	1.06	1.03	0.95	0.87	40.4
Image and Sound Eq.	0.61	0.73	0.65	0.66	1.06	0.89	0.46	77.9
Photography and Photocopy	0.82	0.16	0.55	0.21	0.99	1.42	0.31	104.7
Instruments and Controls	0.81	0.63	0.78	0.79	1.23	1.03	0.45	57.0
Miscellaneous Metal Products	1.12	1.04	0.66	0.87	1.40	0.13	0.33	114.5
Textile & Wood Products	1.84	1.11	1.21	0.22	1.43	1.91	0.69	138.1
Dentistry and Surgery	0.88	1.27	0.65	0.97	1.21	0.75	0.98	36.9
Other	0.38	1.00	0.16	0.74	1.17	0.36	0.27	109.8
Number of FGSIs above unity	*8*	*10*	*5*	*7*	*26*	*10*	*8*	
Specialization (CV_i×100)	*46.2*	*49.9*	*56.6*	*76.6*	*31.7*	*89.7*	*129.7*	

Because no fast growing technologies were identified, *Mining and Wells Machinery and Processes* is excluded from the FGSI computation.

Table 24.A2. RII in Selected Advanced OECD Countries: 1991–2000

Technology	*DE*	*FR*	*UK*	*IT*	*JP*	*NL*	*SE*	$CV_t \times 100$
Inorganic Chemicals	0.87	1.15	1.23	1.02	1.09	1.10	1.61	19.9
Organic Chemicals	1.02	0.88	0.97	0.95	1.06	0.94	1.11	8.0
Agricultural Chemicals	0.98	0.75	1.20	0.67	1.00	0.22	0.17	55.2
Chemical Processes	0.95	0.91	0.97	0.78	1.12	1.06	0.69	16.0
Hydroc. Min. Oils, etc.	1.34	0.88	1.15	1.00	0.89	0.99	0.44	29.2
Bleaching & Dyeing	0.92	0.98	0.74	0.61	0.92	1.67	1.22	34.8
Drugs & Bioengineering	0.96	0.98	1.24	0.86	0.90	0.83	1.02	14.0
Plastic & Rubber Product	0.92	1.04	1.05	0.95	1.01	0.94	0.52	19.8
Materials	0.86	0.87	1.00	0.93	1.06	0.87	1.04	9.1
Food and Tobacco	0.99	0.90	1.10	0.65	1.01	1.03	0.74	18.0
Metal Treatment	0.77	0.66	1.15	0.99	1.13	0.73	0.87	21.6
Apparatus for Chemicals	0.89	0.88	1.10	0.70	1.30	0.94	0.78	21.6
Gen. Non-Elec Ind Eq.	1.02	0.91	0.90	0.92	1.20	1.00	0.68	16.6
Gen. Elec Ind Apparatus	0.80	0.77	1.03	0.72	1.16	0.82	0.83	18.0
Non-Elec Spec Ind Eq.	0.96	1.03	1.11	0.80	1.14	1.00	0.87	12.6
Metal Working Eq.	0.85	0.94	0.91	0.83	1.22	0.92	0.80	15.2
Ass. & Material handling app.	0.85	0.87	0.85	0.93	1.25	0.77	0.67	20.6
Induced Nuclear Reactions	0.89	1.00	0.69	0.50	1.11	0.88	1.47	33.3
Power Plants	0.79	0.75	0.93	0.64	1.39	0.89	0.76	28.0
Road Vehicles and Engines	0.81	0.60	0.81	0.61	1.16	0.76	0.50	28.8
Other Transport Eq.	0.87	0.85	0.70	1.10	1.26	0.83	0.95	20.1
Aircraft	1.03	1.12	0.96	0.65	1.04	1.96	0.89	37.7
Mining & Wells Mach & Proc	0.96	1.19	1.08	1.00	0.74	0.97	1.50	22.1
Telecommunications	0.67	0.83	1.10	0.79	1.09	0.96	1.21	20.6
Semiconductors	0.63	0.89	0.83	0.70	1.10	0.87	0.33	31.6
Elec. Devices and Systems	0.73	0.93	0.97	0.64	1.16	0.96	0.86	19.1
Calculators, Computers, etc.	0.75	0.95	1.13	0.66	1.07	0.99	1.08	19.0
Image and Sound Eq.	0.75	0.96	1.17	0.79	1.06	0.90	0.42	28.4
Photography and Photocopy	0.84	0.44	0.67	0.58	1.05	0.95	0.48	32.7
Instruments and Controls	0.81	0.86	1.01	0.77	1.15	0.91	0.99	14.1
Miscellaneous Metal Products	0.91	0.93	0.81	0.84	1.20	0.79	0.78	16.4
Textile & Wood Products	0.89	1.00	1.12	1.00	1.05	0.86	0.74	13.6
Dentistry and Surgery	1.03	1.10	0.96	0.97	1.06	0.94	0.82	9.1
Other	0.82	0.96	0.90	0.82	1.22	1.24	0.63	23.7
Number of RIIs above unity	*5*	*8*	*17*	*3*	*29*	*6*	*9*	
Specialization ($CV_i \times 100$)	*14.6*	*16.6*	*16.1*	*19.3*	*11.5*	*28.3*	*38.9*	

Chapter 25

USING PATENT CITATION INDICATORS TO MANAGE A STOCK PORTFOLIO

Francis Narin, Anthony Breitzman, and Patrick Thomas

CHI Research, Inc, Haddon Heights, NJ, USA
E-mail:fnarin@chiresearch.com

Abstract: This paper examines the relationship between indicators of technology quality and stock market performance. The purpose of this analysis is to demonstrate how quantitative R&D and technology indicators may be useful tools in the analysis of the stock market. Currently many stock market analysts do not include quantitative technology indicators in their evaluation of companies. The analysis presented in this paper shows how such indicators may be a useful addition to traditional methods of company valuation. The paper describes CHI's technology value model, and presents the results of this model in terms of stock market returns. The results are divided into three sections, according to how often the model is updated. A comparison is made between portfolios updated on an annual, monthly, and weekly basis. The last of these portfolios is based on an actual investment made by CHI using part of its pension fund. The results of the analysis show that updating the model more often than annually improves its performance. This may be owed to the ability to adjust for price changes in stocks during each year.

1.INTRODUCTION

A major objective of science and technology policy analysis is to understand how investing resources in scientific and technological R&D leads to innovations of economic and commercial benefit. There have been numerous studies showing relationships between measures of technological investment, such as R&D expenditures, and outcome measures such as company performance and national and regional technological position. In this paper we shall discuss a direct corporate benefit from investment in technology. The paper reveals a strong association between the quality of

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 553-568.

companies' patent portfolios, as measured by patent citation indicators, and their stock market performance in the short and long-term.

CHI Research, Inc. (CHI) has developed two different, but related, models which use patent indicators to forecast stock price movements. The first model is described in CHI's patent 'Method and Apparatus for Choosing a Stock Portfolio, Based on Patent Indicators' (US Patent 6,175,824), issued in January 2001 to Anthony Breitzman and Francis Narin. This model is based on identifying companies with the strongest patented technology in their industry. Such companies have exhibited consistently strong stock market performance. The second model, developed by Thomas (2001), takes this methodology one step further. It identifies companies which have both strong patent citation indicators and a low stock market valuation. This value approach to technology based investing is discussed in more detail in this paper. Also presented in this paper is a new analysis of the performance of this model based on it being updated monthly, rather than annually.

2. BACKGROUND

In recent years it has become widely accepted that invention and innovation are fundamental forces driving advanced economies. This has led to research showing that the growth in these economies can be traced to the close links between the growth of scientific knowledge and the use of technology (Rosenberg and Birdzell, 1990).

There is also a growing awareness that companies' technology has an important role in the stock selection process. Lev and Zarowin (1998) assert that stock models based purely on financial indicators often fail to produce accurate forecasts of the future performance and stock value of a company. They argue that this is partly because the quality of a company's intellectual capital, as represented by its patent portfolio, trademarks, trade secrets, and other non-financial capital, is becoming an increasingly important element in its performance. Empirical research has lent support to this view. Narin et al. (1987) showed that companies with strong technology had better corporate performance, whilst Deng et al. (1999) showed that strong technology was associated with higher stock market valuations.

Traditionally many stock pricing models have not included technology as a predictive variable. Such models have been developed using various characteristics of company stocks, most of which are based on financial information. For example, O'Shaughnessy (1997) found that stocks with high dividend yields provided higher stock market returns than similar stocks with lower dividend yields. From a value–investing perspective (similar to the perspective used in this paper), Fama and French (1992) showed that market to book ratios could be used to predict future stock price movements. Other researchers have looked beyond company specific issues to examine wider causes of stock price movements, such as economic cycles (McNees, 1992) and specific news events (Cutler et al., 1989). Surveys of the development of stock market analysis can be found in Bhattacharya and Constantinides (1989) and Cuthbertson (1996).

The absence of technology evaluation from many stock price models is caused largely by the inadequacy of public information about firms' R&D activities for the purpose of investment analysis. A firm's periodic R&D expenditures, the sole item related to innovation required to be disclosed in financial statements, is a coarse indicator of the nature, quality, and expected benefits of its science and technology. Firms generally do not disclose information about the nature of their science and technology, and R&D expenditure data does not enable investors to account for differences in companies' innovative efficiency. Furthermore, various innovative activities, particularly in small companies, are often not formally classified as R&D, and are therefore not reported separately to investors. Consequently publicly available information on firms' science and technology is often inadequate for assessing the capabilities of firms to innovate and the impact of such innovations on future corporate performance.

Given the inadequacies of R&D data, many researchers have paid increasing attention to patents as a unit of analysis. The idea of a patent is simple. An inventor or his/her company is granted a twenty year monopoly on an invention, in return for detailed disclosure of how the invention works. This is designed to spur, rather than stifle innovation. The inventor is granted twenty years of exclusive control of his/her invention, whilst the public is able to see how the current invention works, and can therefore build and improve upon the innovation without the pitfalls of starting from scratch.

Patents are becoming increasingly important to commercial organisations, both for internal technological developments, and for generating revenue from licensing initiatives. This has led to a rapid growth in the number of patents issued over the past decade. Given the growth of the patent system and the importance of managing intellectual property, it has become increasingly important to be able to analyse patent portfolios without sifting through thousands of individual patent documents. For this reason, techniques of patent citation analysis have been developed to statistically analyse the quality and strength of patent portfolios.

Patent citation analysis is based on the citations that appear on the front page of patents. When an inventor applies for a patent s/he must show that the invention is novel, useful, and non-obvious to someone with average expertise in the same industry. To do so, the inventor will cite earlier patents, and explain why the new patent improves on the earlier inventions. The patent examiner may also add earlier inventions which limit the scope of the new invention. It is fraud on the patent office not to cite earlier relevant work, and it is also undesirable to cite irrelevant work unnecessarily.

Given that almost all patents cite earlier patents, one can easily count up the citations a patent receives from later patents. The underlying assumption in patent citation analysis is that a highly cited patent (a patent which is referred to by many subsequently issued patents) is likely to contain technological advances of particular importance that has led to numerous subsequent technological improvements. It follows that a company whose patent portfolio contains a large number of highly cited patents is generating high quality technology. Hence one would expect that companies whose patents are highly cited would tend to be more successful innovators, and so perform better in both commercial and capital markets than companies whose patents are cited less frequently.

This does not mean that every important patent is highly cited, or that every highly cited patent is important. However, numerous validation studies have shown the existence of a strong positive relationship between citations and technological importance. Previous research has shown that patents rated highly by an industry's staff were more frequently cited than patents of lower rank (Albert et al., 1991). Recent work also shows that patent renewal and citation frequency are correlated (Thomas, 1999), whilst an early paper showed that patents associated with important inventions were twice as highly cited as control patents (Carpenter et al., 1981). In addition, Breitzman and Narin (1996) showed that pioneering patents are cited five times as frequently as ordinary patents. Prior evidence thus suggests that citations are valid indicators of firms' science and technology.

In addition, economists have in recent years examined the usefulness of patents and patent citations as measures of firms' innovative activities. For example, it has been shown that the intensity of citations to a set of patents is related to the social gains from these patents (Trajtenberg, 1990). In addition it has been shown that patenting is associated with subsequent gains in firms' productivity (Griliches, 1990), and that the intensity of citations to firms' patents is associated with their market values (Hall et al., 1998).

In summary, background research provides a strong rationale for the expectation that companies with strong patent portfolios would perform better in the stock market. Furthermore, information of this type should be particularly valuable because it is not currently available to market analysts, leading to a strong likelihood that the quality of companies' technology might not be properly valued in the market. The following sections detail how CHI, through its technology value model, has demonstrated this link between companies' patent portfolios and their stock market performance.

3. DATA

There are a number of barriers which must be overcome before using patent citations in stock selection models. Perhaps the most complex problem is that of matching patent assignee names to individual companies. Companies may patent under many different names, including subsidiaries and divisional names. It is also a major challenge to account for company mergers, acquisitions, and divestitures. In addition, large numbers of patents are often reassigned from one company to another, many as a result of mergers and acquisitions. Hundreds of thousands of reassigned patents therefore have to be assigned as accurately as possible to the company which currently owns them.

The analysis presented in this paper is based on data extracted from CHI Research's Tech-Line® database. This database contains patent indicators for all organisations which have been issued at least 45 U.S. patents in the previous five years. There are currently around 1,800 of these organisations. CHI has constructed accurate corporate structures for each of these organisations, in order to account for the over 30,000 different assignee names under which they patent.

The subset of organisations used for this analysis contains all U.S. companies listed on the major U.S. stock exchanges (NYSE and NASDAQ) covered by the Tech-Line® database. These companies therefore have at least 45 U.S. patents in the past five years. The model is restricted to U.S. companies listed on U.S. exchanges to remove the effect of any differences between worldwide stock exchanges and currencies. The minimum patent threshold is used to focus the analysis on companies for which patents are an important source of future success. There are currently 450 companies which meet these criteria.

4. METHODOLOGY

The purpose of the value model is to identify companies whose technology is undervalued by the stock market. There are two stages in the modeling procedure. The first stage develops a valuation of companies based on the quality of their technology and their commitment to R&D. In the second stage these valuations are compared with the companies' actual valuations in the stock market. This two-stage process facilitates identification of companies which are under and over valued in the stock market.

In order to place a value on companies based on their patent portfolios, these portfolios were evaluated on an annual basis from 1990 through 1999 using a number of quantitative patent indicators. A comprehensive discussion of these indicators is provided in the background material to the Tech-Line® database (Narin, 1999). The indicators are:

- Number of Patents: The number of patents granted to a company, including its subsidiaries, in the previous year. This is a measure of the technological productivity of a company.
- Patent Growth: The percentage growth in the number of patents granted to a company in the previous year, compared to the year before. This indicator shows trends in a company's commitment to technological innovation.

Current Impact Index (CII): The CII shows the impact of a company's patents on the latest technological developments. It is a measure of how frequently the previous five years of a company's patents are cited by patents issued in the most recent year, relative to all US patents. The CII is a synchronous indicator, and moves with the current year, looking back five years. As a result, when a company's patents from recent years start to drop in impact, this is reflected by a decline in the current year's CII.

Science Linkage (SL): Science Linkage is a measure of the extent to which a company's technology builds upon cutting edge scientific research. It is calculated on the basis of the average number of references on a company's patents to scientific papers, as distinct from references to previous patents. Companies whose patents cite a large number of scientific papers are assumed to be working closely with the latest scientific developments.

Technology Cycle Time (TCT): In general, companies which are innovating rapidly tend to be more successful in product development than companies relying on older technologies. This leads to another citation indicator, the Technology Cycle Time (TCT). TCT is a measure of the median age of the US patents cited on the front page of a company's patents. A tendency to cite older patents is an indication that a company utilises older technology. The average TCT is as short as three or four years in rapidly evolving industries, such as electronics, and as long as fifteen years in industries that change more slowly, such as shipbuilding.

Patent indicators vary greatly across industries. In order to account for these differences the Tech-Line® database divides companies into 26 industry groups, and calculates industry averages for each patent indicator. Industry-normalised indicators are computed by taking the indicator value for a particular company and dividing by the average for that company's industry. By removing the industry effects it is possible to identify the companies which have strong patent indicators relative to other companies in their industry. For example, an automotive company with a Science Linkage of four is more science linked relative to its industry than a biotechnology company with a Science Linkage of eight.

We carried out multiple regression analyses for each year between 1990 and 1998. The independent variables in the regressions were the five raw patent indicators and five industry-normalised patent indicators, along with companies' R&D Intensity (R&D Expenditure/Sales). The dependent variable was the natural log of companies' market to book (MTB) valuation. The MTB measures the relationship between the Market Value of a company (Share Price * Number of Shares Outstanding) and its Book Value (the value of the assets it has on its balance sheet). For example, if a company has a Book Value of $10 million and has 5 million outstanding shares priced at $4 each, it has an MTB of 2 ($20 million / $10 million). The natural log of the MTB was used owing to the skewness of the distribution of MTB values.

The regression analyses revealed that the number of patents and patent growth were not significantly related to companies' market to book ratios. The coefficients associated with these variables were therefore set to zero and the regressions re-run. Setting the coefficient associated with the number of patents to zero means that there is no inherent bias in the model towards large companies with extensive patent portfolios. The model thus depends on the quality of companies' patent portfolios, not their size.

The coefficients for the remaining variables changed each year. However, there was a high degree of consistency across years. The signs of the coefficients were often the same in each year. In most cases there were positive coefficients for CII, SL, and R&D, and a negative coefficient for TCT. However, in a number of years TCT was not significantly related to MTB values.

Owing to the consistency of the coefficients it was possible to combine them to produce a single regression equation that related patent indicators to MTB valuations for the period between 1990 and 1998. The initial coefficients for this equation were the means of the coefficients from the nine models. Sensitivity analysis was then carried out on the equation, with each of the coefficients being changed up to 10% in each direction to establish whether alternative equations would produce values which correlated more closely with actual MTB values. Small changes were made in the coefficients as a result of this analysis. The resultant equation, based on data between 1990 and 1998, was

$$\text{MTB} = e^{(0.4 + 0.4*\text{CIInormed} + 0.15*\text{SL} + 0.011*\text{R\&D} - 0.09*\text{SLnormed})}$$

Two features of this equation are worth noting. The first is the absence of TCT from the equation. This is because of the lack of significance of this variable in a number of years. The second feature is the presence of both SL and SL normed. The former is the raw science linkage indicator, which has a positive coefficient, whilst the latter represents the science linkage normalised by industry, which has a negative coefficient. The influence of the raw science linkage indicator is therefore muted by the negative coefficient of industry normalised science linkage.

The average R2 value of the eight regression equations across the period between 1990 and 1998 was 0.08. F statistics revealed that five out of the eight models were significant at the 1% level, and a further two models were significant at the 5% level. However, the R2 value is relatively low, suggesting that the relationship is a very noisy one. This is a reflection of the complexity of stock market valuation, which leads to a high level of noise in any model of the stock market. For example, Lev and Sougiannis (1996) reported a similar R2 between earnings, a widely used stock market indicator, and MTB.

Substituting a company's patent indicators into the equation above produces an MTB valuation for a company based on its patent indicators and R&D intensity. This valuation is defined as the Technology MTB. Given that the Technology MTB valuations have their foundation in mapping patent indicators against Actual MTB valuations, it might be expected that these two valuations of companies would be similar. However, multiple regression fits a single model to all cases, so that each case has a residual term associated with it. Based upon this residual it is possible to define whether the company, based upon its technology, is overvalued (Actual MTB > Technology MTB) or undervalued (Technology MTB > Actual MTB).

Companies with the largest relative residuals are of particular interest. These are the most undervalued and overvalued companies in the sample. To identify these companies, all companies in the sample were placed in percentiles according to their Technology MTB, with 100 assigned to the company with the highest Technology MTB, and 1 to the company with the lowest Technology MTB. Companies were then placed into percentiles according to their Actual MTB, with 100 representing the highest Actual MTB.

For each company the Actual MTB Percentile was subtracted from the Technology MTB Percentile. Companies were then placed into percentiles on the basis of the resultant differential, to produce the Investment Potential of each company. The highest Investment Potential (100) was assigned to the companies with the largest positive differential. The Technology MTB Percentile of these companies exceeded their Actual MTB Percentile by the largest amount. These companies were therefore the most undervalued in the sample. The lowest Investment Potential (1) was assigned to the most overvalued companies in the sample, whose Actual MTB Percentile exceeded their Technology MTB Percentile by the largest amount.

The Investment Potential of a company reflects how its valuation in the stock market compares with a valuation of it based solely upon its technology. A company with an Investment Potential of 100 has strong technology that is not recognised by the stock market. Meanwhile, a company with an Investment Potential of 1 has a valuation in the market that cannot be justified on the basis of its technology (although there may be other factors that explain its high valuation). A company with an Investment Potential of 50 is regarded as fairly valued, based upon the quality of its technology.

5. PERFORMANCE OF THE VALUE MODEL UPDATED ANNUALLY

The usefulness of a model of this type depends upon its ability to forecast future changes in stock prices. To measure this we paid particular attention to the companies at the two ends of the Investment Potential distribution. These are the companies which are most under and over valued. Figure 25.1 shows the returns that investors would have received if they had invested equal dollar amounts in two annually updated portfolios containing the twenty most undervalued companies and the twenty most overvalued companies.

Each year the portfolio is sold, and the funds invested in the new portfolio of stocks. Figure 25.1 reveals that an investment of \$100 in December 1990 in an annually updated portfolio of the twenty most undervalued stocks would have returned over \$2,100 by December 2002.

This portfolio significantly outperformed both the NASDAQ Composite Index and the S&P 500. Meanwhile, the portfolio of overvalued stocks returned only \$300 over the same period, underperforming both the NASDAQ and the S&P 500. From 1999 onwards, when the model was no longer based on back data, but was being updated in real time, the performance continued to be very strong.

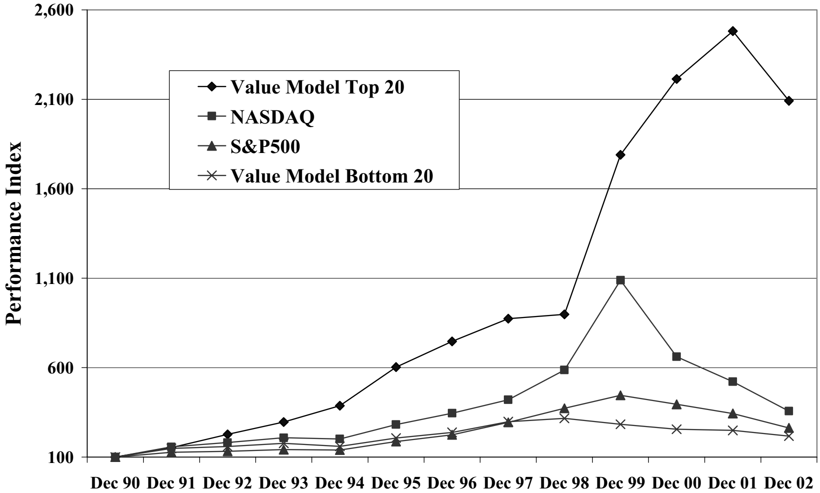


Figure 25.1. Performance of the value model 1991–2002

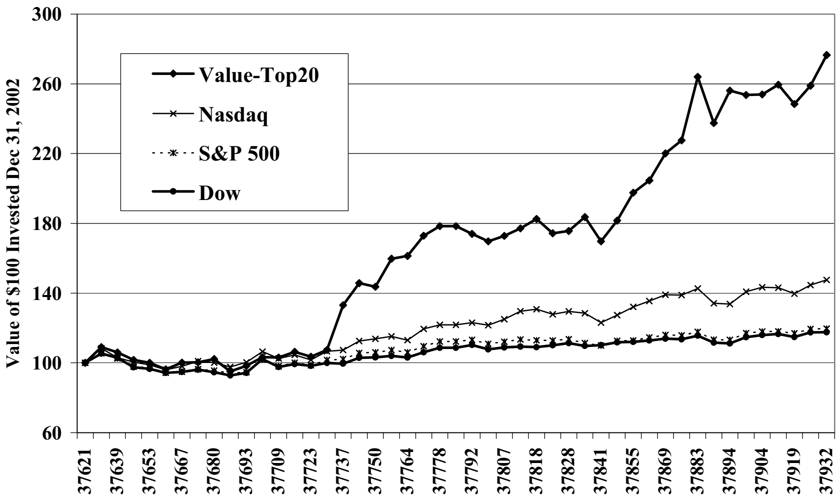


Figure 25.2. Value model performance from Dec 31, 2002 to Nov 7, 2003 (based on investment of $100 on Dec 31, 2002)

Although 2001 and 2002 were very weak years for the markets as a whole, the twenty most undervalued stocks only decreased about 4%, whilst twenty most overvalued stocks decreased 15%. Over the same period the NASDAQ and S&P 500 each decreased by between 40% and 50 %.

Figure 25.2 shows the performance of the value model from the beginning of 2003 through to the 7th of November 2003, the date on which this analysis was implemented.

Also shown on the chart are the performance figures for the Dow Jones Industrial Average, the S&P 500, and the NASDAQ.

The stock market has followed an upward trend in 2003, with both the Dow Jones Industrials and the S&P 500 up approximately 20%, and the NASDAQ up more than 40%. Even in this upward market the performance of the value model has been exceptionally strong. The twenty most undervalued companies, selected at the close of the markets on December 31, 2002, are up 187% over this ten month period. As shown in Table 25.1, all of these twenty stocks gained value during this time.

Table 25.1. Value on 7th November 2003 of $50 Investments in Each of the 20 Top Ranked Companies at Close of Business on 31st December 2002

Ticker	*Shares in Portfolio*	*Share Price 31st Dec 2002*	*Share Price 7th Nov 2003*	*Value 7th Nov, 2003*	*Percent Gain*
CALP	16.89	2.96	6.39	107.94	115.88
NERX	116.28	0.43	6.06	704.65	1309.30
EMIS	14.37	3.48	7.90	113.51	127.01
IMMR	42.74	1.17	6.00	256.41	412.82
CIEN	9.73	5.14	6.90	67.12	34.24
BRKS	4.36	11.46	26.47	115.49	130.98
MTIX	23.26	2.15	3.65	84.88	69.77
GLFD	12.56	3.98	6.71	84.30	68.59
JDSU	20.24	2.47	3.49	70.65	41.30
AMKR	10.50	4.76	19.48	204.62	309.24
CNXT	38.17	1.31	5.69	217.18	334.35
ARDM	30.86	1.62	2.14	66.05	32.10
COHR	2.51	19.96	23.93	59.94	19.89
NOVL	14.97	3.34	7.50	112.28	124.55
GLW	15.11	3.31	12.00	181.27	262.54
COMS	10.80	4.63	7.44	80.35	60.69
ACO	8.62	5.80	15.35	132.33	164.66
FON	3.45	14.48	15.24	52.62	5.25
MLIN	15.63	3.20	5.26	82.19	64.38
LYNX	17.06	2.93	4.99	85.15	70.31
				Average Return	**187.89**

One stock, NERX, which rose from 43 cents per share at the start of the year to $6.06 on November 7th, accounts for 38% of the portfolio's gain. However, nine other stocks doubled in value over the 10 months, so the performance is not owed simply to one outstanding performer. In fact, if NERX were removed, a portfolio containing the other 19 stocks would still have more than doubled in value over the ten month period.

6. PERFORMANCE OF THE VALUE MODEL UPDATED MONTHLY

Updating the value model annually has resulted in strong performance over recent years. However, few portfolio managers buy stocks and then hold them for a year, without adjusting for changes in the market and in the stocks themselves. In the value model the relationship between companies' Technology MTB and Actual MTB changes constantly with changes in the price of their stock. The performance resulting from updating the model on a monthly, rather than annual, basis is therefore of great interest in real portfolio management.

We carried out a brief analysis comparing the monthly model for 2002 with the annual model. Simultaneously, we looked at different portfolio sizes for the monthly model, specifically at models based on the top 10, 20, 30, 40, and 50 companies, as ranked at the beginning of each month. Figure 25.3 shows the performance of these portfolios, along with the performance of the NASDAQ (QQQ), S&P 500, and the top twenty stocks from the annual model.

The worst performance in Figure 25.3 is that of the NASDAQ, which was down more than 35% in 2002, followed by the S&P 500, which fell more than 20%. Over the same period the top twenty stocks from the annual model lost 15%. In sharp contrast to these figures the monthly model's top ten stocks increased more than 10% in value, and the top twenty and top thirty also provided positive returns. However, by the time forty stocks were included in the monthly value portfolio the portfolio had an overall loss for the year.

Clearly, in 2002 an approach in which the portfolio is refreshed at the beginning of each month would have performed significantly better than a portfolio bought and held for the whole year. A similar analysis comparing the monthly model with the annual model for 2003 yields similar results, with the monthly model performing even better than the annual model by November 7th.

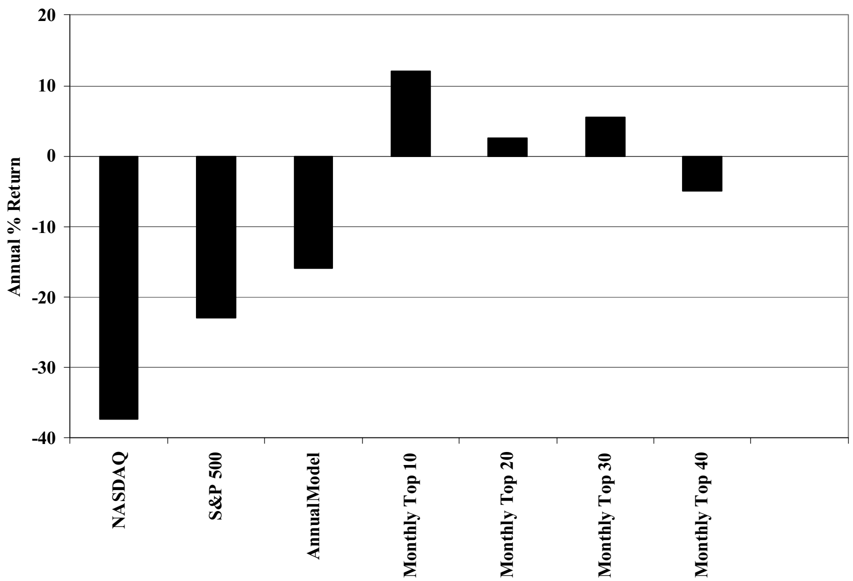


Figure 25.3. Year 2002 Comparative Performance- Annual and Monthly Models, plus NASDAQ and S&P500

This raises the question of why a portfolio updated monthly should exhibit such strong performance, especially because citations play a major role in the model, and these citations reflect technology from earlier years. The result may be explained by referring to the two-stage process involved in the model. The first stage, in which companies' technology is evaluated, produces relatively consistent results over time. Patent indicators tend to change gradually, so the difference between an annual model and a monthly model is relatively small in terms of the first stage of the model.

In the second stage of the model, however, stock market valuations are introduced. These valuations can change markedly over a short period of time. If the model is updated annually, changes in companies' stock prices during the year are not taken into account. Updating the model on a monthly basis enables these changes to be incorporated. For example, a company may experience a sharp decline in its stock price during the year. As a result it becomes undervalued when based on its technology strengths.

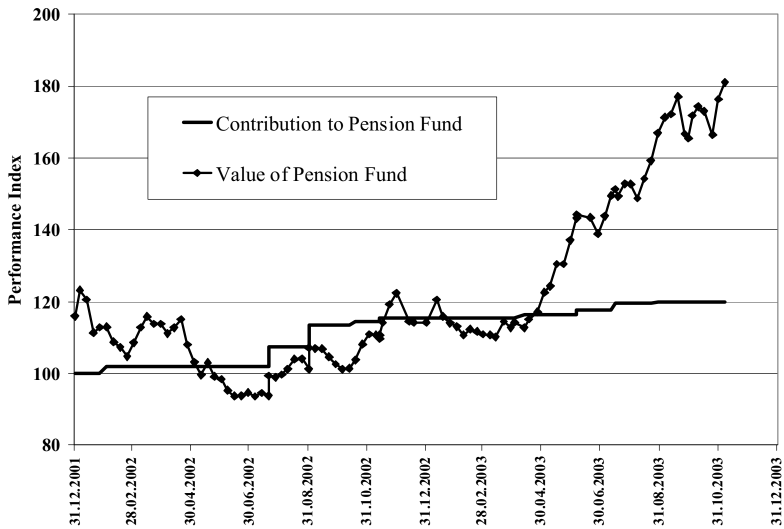


Figure 25.4. Actual Pension Performance Jan 1, 2002 to Nov 7, 2003 (Based on Index where Jan 1, 2002=100)

Using the monthly model this company can be introduced to the portfolio during the year. It could not be introduced using the annual model, since the portfolio is set at the start of the year.

7. PERFORMANCE OF CHI'S REAL PENSION PORTFOLIO

In July 1999 CHI decided to perform a real life test of our methodology. We invested a small portion of the CHI employee pension fund in a portfolio based on the top ranked stocks in the value model. We still manage a small percentage of the pension fund using these rankings, along with standard portfolio management tools. Below we report the performance of this portfolio, in terms of average gain per week.

In latter part of 1999 the pension portfolio did very well, averaging about 0.7% gain per week in the strong stock market of that period. The performance remained in the range of 0.5% to 0.7% per week until 2001, when it was negatively affected by the market crash after September 11th, 2001. By the end of October of 2002 the average weekly gain since inception was 0.19% per week.

Since October 2002 the portfolio has enjoyed a strong ascent, as shown in Figure 25.4. This figure shows the money contributed to this part of the pension fund, and also the value of this fund over time. The increase in the value of the portfolio over this period coincided with an upturn of the stock market, particularly in technology stocks. Correspondingly, by the beginning of November 2003 the average weekly gain from the time we started the portfolio in mid-1999 is 0.43% per week. This is a high return given the volatile stock market conditions over this period.

8. CONCLUSIONS

In this paper we have shown a strong association between the quality of companies' patent portfolios, as measured by patent citation indicators, and their stock market performance. This demonstrates a direct benefit for companies investing in high quality research and development efforts. A particularly interesting aspect of this finding is that the number of patents a company holds is not a significant predictor of its performance. The important factor is the quality of a company's patents, rather than their number.

REFERENCES

Albert, M., Avery, D., McAllister, P., Narin F. (1991). Direct validation of citation counts as indicators of industrially important patents. *Research Policy*, 20, 251–259.

Bhattacharya, S., Constantinides, G. (Eds.). (1989). *Theory of valuation: frontiers of modern financial theory*, Vol. 1. Totowa: Rowman & Littlefield.

Breitzman, A., Narin, F. (1996). A case for patent citation analysis in litigation. *Law Works*, 3, 3, March.

Carpenter, M., Narin, F., Woolf, P. (1981). Citation rates to technologically important patents. *World Patent Information*, 4, 160–163.

Cuthbertson, K. (1996). *Quantitative financial economics: stocks, bonds and foreign exchange*. Chichester: John Wiley & Sons.

Cutler, D., Poterba, J., Summers, L. (1989). What moves stock prices. *Journal of Portfolio Management*, 15 (3), 4–12.

Deng, Z., Lev, B., Narin, F. (1999). Science & technology as predictors of stock performance. *Financial Analysts Journal*, 55, 3, 20–32.

Fama, E., French, K. (1992). The cross section of expected stock returns. *Journal of Finance*, 47, 427–466.

Griliches, Z. (1990). Patent statistics as economic indicators: A Survey. *Journal of Economic Literature*, 28, 1661–1707.

Hall, B., Jaffe, A., Trajtenberg, M. (1998). *Market value and patent citations: a first look.* Paper prepared for the Conference on Intangibles and Capital Markets, New York University.

Lev, B., Sougiannis, T. (1996). The capitalization, amortization, and value–relevance of R&D. *Journal of Accounting and Economics*, 21, 107–138.

Lev B., Zarowin, P. (1998). *The boundaries or financial reporting and how to extend them*. Presented at The Conference on Intangibles and Capital Markets, New York University, May 13.

McNees, S. (1992). How large are economic forecast errors, *New England Economic Review*, July/August, 25–42.

Narin, F. (1999). *Tech-Line® background paper*. In J. Tidd (Ed.), Measuring strategic competence. Imperial College Press, Technology Management Series.

Narin, F., Hamilton, K., Olivastro, D. (1997). The increasing linkage between U.S. technology and public science. *Research Policy*, 26 (3), 317–330.

Narin, F., Noma, E., Perry R. (1987). Patents as indicators of corporate technological strength. *Research Policy*, 16, 143–155.

O'Shaughnessy, J. (1997). *What works on Wall Street*. New York: McGraw-Hill.

Rosenberg, N., Birdzell L, Jr. (1990). Science, technology and the Western Miracle. *Scientific American*, 263 (5), 42–54.

Thomas, P. (2001). A relationship between technology indicators and Stock market performance. *Scientometrics*, 51 (1), 319–333.

Thomas, P. (1999). The effect of technological impact upon patent renewal decisions. *Technology Analysis & Strategic Management*, 11 (2), 181–1997.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *Rand Journal of Economics*, 21, 172–187.

Chapter 26

PATENT DATA FOR MONITORING S&T PORTFOLIOS

Koenraad Debackere[1] and Marc Luwel[2]

[1] K.U.Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, Steunpunt O&O Statistieken, Leuven, Belgium. E-mail: Koenraad.Debackere@econ.kuleuven.ac.be
[2] AWI, Ministerie van de Vlaamse Gemeenschap, Brussel, Belgium

Abstract: This chapter deals with the use of patent data to monitor science and technology (S&T) portfolios. S&T portfolios have become central tools for examining and for monitoring the vitality of institutions, innovative clusters, and regions in the innovation game that underpins their respective economic growth and development. Those portfolios have to be monitored not only at the intra-organisational level, but also at the inter-organizational level, as well as at other appropriate levels of analysis for designated systems of innovation (e.g., specific technology clusters). To this end, the development of appropriate, easy to use and transparent, benchmark indicators to assess the relative strengths and weaknesses of S&T portfolios is important. In this chapter the construction of a particular type of benchmark indicator, based on relative specialization indices, is reported and its usefulness is assessed by its application to the European Patent Database.

1. MONITORING S&T PORTFOLIOS

Portfolio management in science and technology is not new. Ever since the development of the concept of technological S-curves many years ago (see for example: Martino, 1983; Girifalco, 1991; Porter et al., 1991; Roussel et al., 1991; or Floyd, 1997), companies have developed methods to monitor and to assess the potential and the relative quality of their science and technology investments. The concept of S-curves pointed to the explicit risks and uncertainties involved in developing new technological capabilities and applying them towards the fulfilment of product–market needs. They also provided an attempt to extrapolate the speed at which new technological trajectories would diffuse and become common technological practice

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 569-585.

(Sahal, 1981). As most companies manage a myriad of projects attempting at major as well as minor improvements of their current Science and Technology (further abbreviated as S&T) base, it became obvious that S-curves were just one criterion relevant to assessing the vitality of a corporate S&T portfolio. Risk–reward criteria, as well as indicators of competitive dynamics such as a company's S&T position versus those of competitors, became standard concepts. Those analyses showed that not all S&T endeavours could be considered equal. Some were indeed more fundamental than others. Abernathy and Clark (1985) were amongst the first to discern different types of S&T efforts within a company. Some of those efforts would indeed disrupt the technological competences of the company, whilst others would just enhance those competences in a somewhat incremental way. Along a second dimension they stated that a company's S&T efforts might either destroy or enhance existing market and distribution relationships.

Combining the market and technology dimensions, they constructed a two by two-dimensional model assessing the transilience, or impact, of various types of S&T efforts. They coined them: regular (enhancing both the existing technology and market competence of the company); niche (enhancing the existing technology competence but destroying the market competence); revolutionary (destroying the technology competence, but enhancing the market competence) and finally, architectural (destroying both the existing technology and market competence of the company). The resulting 'transilience map' proved to be an interesting tool to map and to assess a company's S&T portfolio. The central units of analysis in this assessment became the types of product–related S&T projects a company was undertaking in its R&D departments.

The 'transilience map', which was first published by Abernathy and Clark in 1985, was characteristic of the onset of a wide array of research efforts aimed at understanding and developing methods and tools for assessing and managing the multiple S&T projects going on within a company. As project management techniques no longer sufficed, multi-project management techniques were developed. The S&T portfolio became both the method and the tool to handle the complexity of this multi-project environment (e.g., Roussel et al., 1991;, Wheelwright and Clark, 1992; Floyd, 1997; Meyer and Lehnerd, 1997; and Cooper et al., 1997a&b). Typical S&T portfolio management at the company level includes assessing and mapping the following dimensions of the portfolio:

– The degree of technological maturity of the various S&T projects in the portfolio (typically according to such notions as 'embryonic', 'growing', and 'mature', as described in Foster (1986) or Roussel et al. (1991));

- The market and financial 'attractiveness' of the S&T projects being proposed or executed (Wheelwright and Clark, 1992; Roussel et al., 1992; Brown and Eisenhardt, 1995);
- The risks incurred with the various S&T projects, typically along such dimensions as technological risks, financial risks, commercial risks, and, increasingly, operational risks (see Roussel et al., 1991);
- The potential rewards of the various S&T projects, using such standard financial techniques as Net Present Value calculations or real options modelling (Jägle, 1999; Perlitz et al., 1999; Angelis, 2000; Boer, 2000; McGrath and MacMillan, 2000);
- The competitive position, focussing on strengths and weaknesses, the company has achieved in the various S&T projects proposed or selected viz. its main competitors (Meyer and Lehnerd, 1997; Cooper et al., 1997a&b; Cooper ,2001);
- The presence or lack of competences with respect to the definition, the implementation and the timely execution of the various S&T projects in the portfolio (Nonaka and Takeuchi, 1995; Bone and Saxon, 2000; Cooper et al., 2000).

The next step in these portfolio management approaches typically is an analytical one. The various criteria just listed are subjected to both univariate and multivariate analyses. The multivariate analyses allow for screening the variance and the covariance within the portfolio on the different dimensions and criteria utilised. A typical example of a portfolio with five S&T projects is presented in Figure 26.1. This map shows the distribution of the five projects along two dimensions: probability of project success (x-axis); and financial return as measured via an NPV-calculation (y-axis). The sizes of the bubbles in the bubble chart represent the magnitude of the respective project budgets. Typically such maps are then subjected to various analyses and interpretations.

The univariate analysis will usually list and rank the various projects according to their absolute scores on the different criteria used. It is a simple, first–order statistical frequency analysis. The second step will then be to look at the variance accross the different projects. Here decision makers want to address such questions as: what is the risk profile we are willing to tolerate for our company given a portfolio of projects? Or, what is the technology specialization profile we want to acheive at our company? Finally, one is not only interested in distributions and variances, but also in correlation and covariance. In other words: to what extent are the different projects independent of one another along the various dimensions that have been used to analyze the portfolio. Do there exist important spillovers between the various projects, or not? Spillovers can be determined in terms of technical spillovers as well as resource or market spillovers.

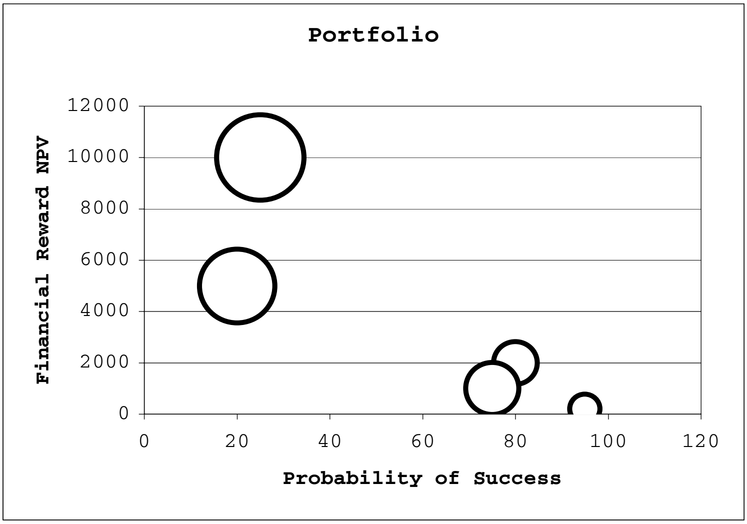


Figure 26.1. Portfolio Map with Five S&T Projects

Finally, the criteria and their analysis are embedded in a decision making framework that attempts at synthesis. In other words, the end result should be a selection of those S&T projects that can either sustain or rejuvenate the company's market position. This synthesis is typically the outcome of a triangulation process which balances (1) the attractiveness of the individual projects against (2) the spillovers and interproject synergies to be obtained, taking into account (3) the resource availability at the company.

In Figure 26.2 a graphical representation of this decision synthesis is provided. It points to portfolio selection, in the end, boiling down to an iterative process of triangulating and balancing the three cornerstones just described. This is an exercise requiring both top-down and bottom-up interactions. The top-down interactions are needed to legitimate and to institutionalize momentum. The bottom-up interactions are required to create and to build momentum.

2. STRETCHING THE S&T PORTFOLIO BEYOND COMPANY BOUNDARIES

The concept of S&T portfolios as described above need not be confined to intra-company decision making. Portfolio assessment and mapping can happen at other levels of analysis as well. More specifically, the tool can be useful for monitoring and for assessing the performance of organizations and clusters of organizations (e.g., the biotechnology cluster in a particular region) which operate within specific 'systems of innovation'. Systems of innovation have been defined at various levels of analysis.

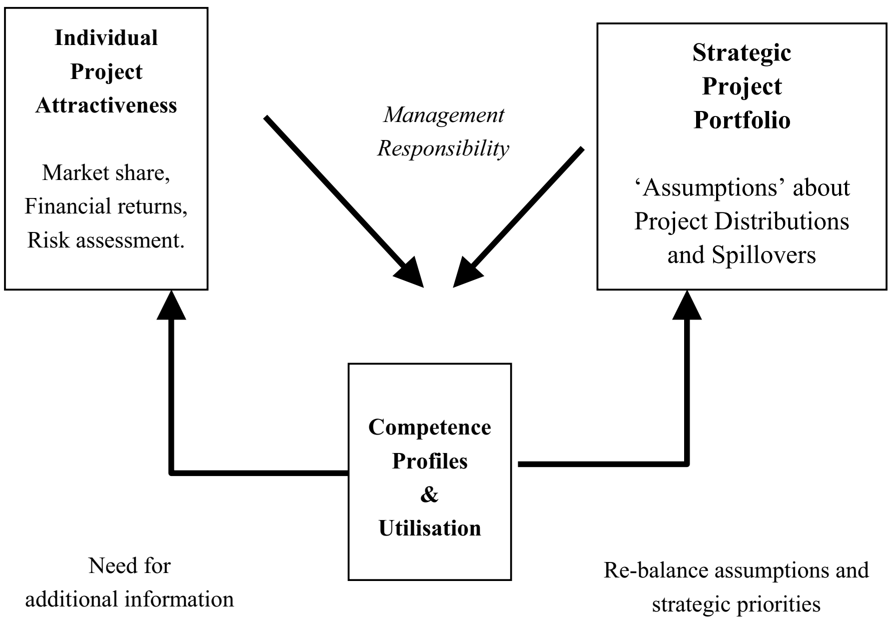


Figure 26.2. The S&T Project Portfolio Decision Making Process

Recently, regional as well as national systems of innovation have received ample attention (Dosi et al., 1989; Antonelli, 1995) when studying the dynamics of economic growth and development. For example, at the regional level the concept of technology clusters (combining the presence of public research institutes and the R&D efforts of private companies) emerged as a relevant and interesting unit of analysis. At this level of analysis one is, of course, not so much interested in the risk profile of a portfolio of projects. Rather, what one might want to assess, amongst various other issues which warrant attention, are the strengths and weaknesses in the relative technology specialization of the R&D organizations (public or private) within the innovation system to be studied.

A system of innovation is the set of supportive arrangements, actors, and their interactions that accounts for the innovation potential and capability of a region or nation. It is obvious that systems of innovation can be compared with one another. The European RITTS projects have offered opportunities for benchmarking the strengths and weaknesses of regional innovation systems (Nauwelaers, 2000). Forty European regions have been studied as to their innovation potential and achievements. The roles and the contributions of various actors in each of the regions have been assessed and documented.

Input indicators, such as R&D personnel and R&D expenditures, have for figured on the agenda of many regional and national comparisons. However, output–oriented indicators such as publications and patents (see: Debackere et al., 1999; or Luwel et al., 1999) may have received still more attention. Output–oriented indicators are well suited to assisting in monitoring the strength and the vitality of a country's or region's S&T portfolio. Just as a company's S&T portfolio allows it to benchmark the strengths and weaknesses of its various S&T projects, so can a region's S&T portfolio allow for a comparison of the strengths and weaknesses of its various S&T actors. As a consequence, substituting actors for projects enables one to stretch the boundaries of portfolio utilization from companies to other units and levels of analysis, such as for instance, the various actors in a regional innovation system. This is the aim of the remainder of this chapter: applying the concept of S&T portfolios to develop a possible benchmark indicator on the technological vitality or fitness of the various actors within a regional system of innovation.

When applying the concept of an S&T portfolio to this level of analysis it is important to design a transparent and consistent set of indicators which are robust and which allow for the straightforward replication across various levels of analysis relevant to a regional system of innovation. Starting from the European Patent Database and applying the concept of the Relative Specialization Index (Balassa, 1961), we have developed a portfolio mapping technique for studying the relative technological specialization of companies in a region (in specific fields of technological activity) versus relevant international control groups.

3. PATENTS AS A SOURCE OF DATA TO BENCHMARK S&T PORTFOLIOS

Patent data have been widely used in many innovation studies (Griliches, 1984; 1990; Schmoch et al., 1992). Next to patent count data, it is obvious that patent documents, because of the legal 'reporting' requirements, provide the researcher with a wealth of information which can be used for various types of research questions and analyses. For instance, typical patent documents contain the names and the addresses of the inventors and their applicants, as well as references to other scientific and technological documents. This information can be easily used to map progress and collaboration in technological fields as well as to assess the vitality of various organizations (firms as well as universities) in a particular field of technological development or in a particular system of innovation. Scholars such as Francis Narin (1987, 1988 & 1997) have been extremely prolific in using patent data as a source of data yielding insights beyond the 'mere' number counts and citation analyses. Two major sources of patent data are the *European Patent Office* (EPO) databases and the databases by the U.S. Patent and Trademark Office (USPTO).

Compared to the USPTO data, EPO data allow for disentangling in detail patent applications and patent grants. Indeed, in the U.S. system patents were (until 2000) only listed in the USPTO databases once they had been granted. In the European system this is not the case. Eighteen months after filing the patent the full document is disclosed, regardless of whether it has been granted or not. The USPTO system has recently also moved in that direction.

Of course, not all patents filed are eventually granted. There are two major reasons for this difference. The *first* one is obvious. Whenever the patent request does not live up to the expectations of newness, inventiveness, and enablement, the patent will not be granted.

A *second* explanation is more strategic in nature. We have already discussed the rising importance of patent portfolios in global market competition (Debackere et al., 1999). Just as patent portfolios may impede entry into specific product markets and curtail international expansion strategies of competitors, filing for patents without having the intention of pursuing the complete patent application trajectory may be part of a pre-emptive strategy. Indeed, when filing for a European patent the applicant knows in advance that the application will be published eighteen months later, and hence from that point in time onwards, belong to the public domain. By doing so, the applicant may intentionally pre-empt others from staking claims to a similar invention.

Since patents differ greatly in quality (see for instance Trajtenberg, 1990), scholars have long since sought to assess the value of individual patents. Three approaches have been subject to extensive research and have acquired a status of being valid measures when it comes to assessing patent quality. They are: (1) the patterns of citation to specific patents; (2) the extent to which patent renewal fees are paid; and (3) the geographical scope of the patent protection requested. In this respect the lack of citation information in the regular EPO data is unfortunate. The existence of the REFI database, which lists the references cited in the prior art search reports, can remedy this lack of information in the regular EPO databases to a certain extent.

For the construction of a transparent and easily used benchmark portfolio map, only patent count data are used. Both patent applications and patent grants have been considered. Patent applications are considered to be closer to the input side of technology creation (serving as a proxy measure of the creation of new technologies). Patent grants are considered to be closer towards the output end of the technology creation process (thus serving as a proxy for the exploitation of results of technological creativity).

In a total of about 750,000 patent applications available in the volume 1997/001 of Espace Bulletin, covering the period December 1978 December 1996, 9,537 patent applications have a Belgian applicant and/or inventor. Patent data have been assigned to the different Belgian regions on the basis of the addresses of the applicants and/or inventors. Given our aim of benchmarking regional S&T positions, this was a necessary step in our analysis. Belgium consists of three different regions: Flanders; Wallonia; and Brussels. Flanders located in the North of Belgium is the largest region, representing about 60% of Belgian GDP. Slightly over 67% of all Belgian patent applications have a Flemish applicant and/or inventor. On average about 47% of all EPO patents applied for are eventually granted. This average holds for the Belgian case as well as for the total EPO database (Vlaams Indicatorenboek WTI, 2003).

The patent database was further extended with additional layers of data. Patent data are connected to economic data, in order to further assess the technological and the economic position of Belgium and Flanders. These data layers included VAT data on production statistics and export statistics, as well as data on the structure of the companies holding the patents (independent or part of multinational corporate structures). Previous analyses (reported in Debackere et al., 1999) have pointed to the overwhelming importance and presence of twenty companies in the total Belgian and Flemish patent portfolio. These companies, which account for about 63% of all Flemish EPO patents, will be used as the empirical basis for the development of the specific benchmark methodology based on a technology specialization profile analysis.

4. CONSTRUCTION OF A PATENT–BASED S&T PORTFOLIO BENCHMARK

In order to develop the benchmark method, we use a 'Relative Specialization' measure as first developed by Balassa (1961), but which is now adapted to measure the Relative Specialization Index (RSI) of organizational entities in specific technological areas. The technological areas are derived using the IPC codes as a classification scheme.

$$\mathbf{RSI_{ij}}$$

(Relative Specialization Index of Organization j in Technological Area i)

$$= \mathbf{(P_{ij}/P_j) \,/\, (P_i/P)}$$

with

P_{ij}: number of patents of organization j in area i,
P_J: number of patents of organization j in all areas,
P_i: number of patents of group of organizations studied in area i,
P: number of patents of group of organizations studied in all areas.

RSI_{ij} compares the share of EPO patents held by an organizational entity in a certain technology area (operationalised via IPC codes or IPC code clusters such as the ones developed by the Fraunhofer Institute) with the similar share of the group considered in the benchmarking exercise. We now apply this index in the following manner. Given the nature of the Balassa Index, we can only apply the method described hereafter in those instances in which there are sufficient numbers of patents per clustering cell and per organization considered in the analysis. Hence the benchmark method developed in this chapter only applies to organizations having developed sufficient levels of patenting activity, otherwise one might obtain trivial results. Indeed, a company 'not specialized' in a particular area may be trivially 'strong' in this area in terms of mere patent number.

Step 1. For every company or organizational entity (further referred to as the 'target company') that needs to be benchmarked in the comparison group, we calculate the following weighted Relative Specialization Index for the complete portfolio of IPC domains in which the company is active with a sufficient patenting activity in terms of numbers of patents in those domains (> 50) in order to avoid 'trivial' results as mentioned above:

$$\mathbf{PSI_j}$$
$$\textbf{(Portfolio Specialization Index of Target Company j)}$$
$$=$$
$$\mathbf{\sum_i w_i\, RSI_i}$$

with i = 1 … N the number of IPC classes in which target company j is active,

RSI_i the Relative Technological Specialization Index of target company j in IPC class i (see formula described above),

w_i the relative weight of IPC class i in the total patent portfolio of target company j, thus w_i is the fraction of the patent total of j in IPC class i.

Step 2. For the purpose of the analyses reported in the construction of the benchmark portfolio map, we have selected comparison groups of relevant EPO companies having sufficient numbers of patents in the relevant IPC classes against which the target companies are to be compared.

Once this benchmark group has been constructed we then calculate the following indices. For each of the benchmark EPO companies, we first select all those IPC classes that overlap with the IPC classes in which the target company is active. The summed set of patents in the overlapping IPC classes now becomes a central measure for further comparisons between the target company and the relevant control group. Each benchmark company has a Relative Specialization Index for each of the overlapping IPC classes (taking into account, once again, the caveat that we need sufficient numbers of patents per class; hence only entities that have sufficient levels of patenting activity per and across the IPC classes considered in the benchmark analysis, can and should be compared with one another using this method).

However, the benchmark companies can also be very active in IPC classes which differ from the ones that overlap with the target company. In other words, there exist overlap and non-overlap IPC classes for the benchmark companies. In addition, there exist benchmark companies which do not show any overlap, but which nevertheless have developed strong positions in other IPC classes. We can demonstrate this situation as follows.

Assume the target company is highly active across five different IPC classes:

AD	**BC**	**GG**	**HY**	**KL**

Assume a chosen benchmark EPO company is active in 12 IPC classes, five of which are overlap classes with the target company:

AA	**AB**	**AD**	**BC**	**GG**	**HY**	**KL**	**MM**	**PO**	**VG**	**WS**	**YH**

Based on this IPC class sequencing information we now compute two new Portfolio Specialization Indices for each company in the benchmark group of companies. They are called the Overlap Portfolio Specialization Index and the Portfolio Specialization Index. They are defined in the following manner.

$$\mathbf{OPSI_j}$$

(Overlap Portfolio Specialization Index of Company j)

$$= \sum_k \mathbf{w_k}\ \mathbf{RSI_k}$$

with k = 1 ... N the number of IPC classes in which company j overlaps with the IPC classes of the target company,

RSI_k the Relative Technological Specialization Index of company j in IPC class k,

w_k the relative weight of IPC class k in the overlapping part of the patent portfolio of company j with the target company. Thus w_k is the fraction of the total number of patents in IPC class k, viz., the total number of patents across all overlapping IPC classes in the comparison with the target company.

(Hence if company j has a total of 1000 patents in its portfolio, of which 300 in overlapping IPC classes, then the weight will be fractionated against the denominator of 300 and NOT of 1000 during the calculation of the Overlap Portfolio Specialization Index).

Thus the Overlap Portfolio Specialization Index is a weighted specialization index, showing the relative position of the benchmark companies, viz., the target company, but limited to those IPC classes or technological domains which the target company has developed in its portfolio.

In other words, whenever the OPSI value of a benchmark company is lower than the PSI of the target company, it means that the benchmark company is less specialized than the target, at least in the technological domains of the target. If, on the other hand, the OPSI value of a benchmark company is higher than the PSI value of the target, then this points to a relative advantage of the benchmark company over the target company. As a consequence the OPSI–PSI comparison allows for an analysis of the relative strengths and weaknesses of the selected target company, viz., a relevant group of benchmarks.

These benchmarks can be chosen as a function of the analysis focus deemed relevant within a particular system of innovation, e.g., assessing the relative position of the biotechnology cluster in a particular region. Of course, when making the selection we also have to take into account the computational requirements of the Balassa Index in order to avoid trivial results. Hence the target companies and entities are to be carefully chosen by the researcher or analyst, depending upon the portfolio analysis she or he intends to conduct.

As stated, however, the benchmark companies or entities can and will also be active in IPC classes which differ from those overlapping with the target company or entity. Hence the need to compute a second Portfolio Specialization Index for each of the benchmark companies or entities. This second Index simply is the total weighted Portfolio Specialization Index computed across all IPC–classes in which a benchmark company is active.

$$\mathbf{PSI_j}$$

(Portfolio Specialization Index of Benchmark Company j)

$$= \sum_k w_k \, \mathrm{RSI}_k$$

with k = 1 … N the number of IPC classes in which company j is active,

RSI_k the Relative Technological Specialization Index of company j in IPC class k,

w_k the relative weight of IPC class k in the total patent portfolio of company j.

Step 3. Using these computations, the following positioning map can now be derived. For each company, transformed PSI and OPSI indices can be computed according to the formula:

$$\mathbf{100 \text{ x } (Index^2\text{-}1/Index^2\text{+}1)}$$

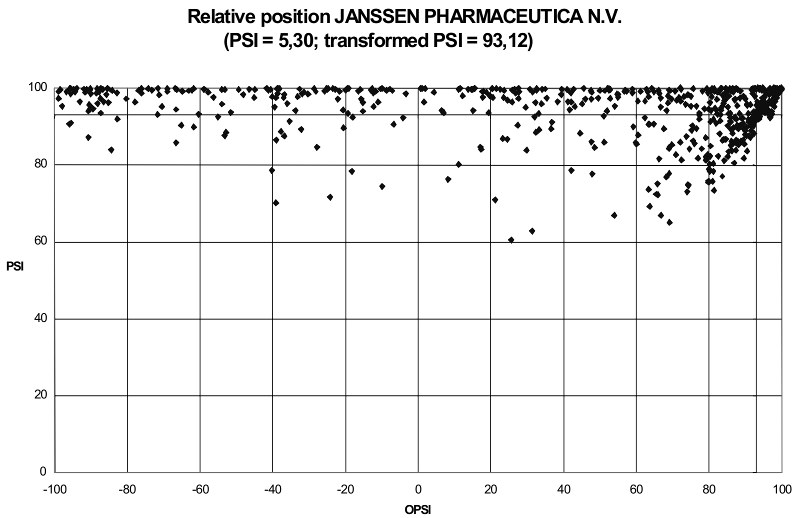
This analytical step results in four quadrants across which the group of companies or relevant units of analysis is distributed, viz., a particular target company or unit of analysis. The quadrant W/W combines all companies which are less specialized than the target company, as well for the overlapping part of their technology portfolio as for their total technology portfolio. The quadrant S/S combines all companies which outperform the target company, as well for the overlapping part of their technology portfolio as for their total technology portfolio. The quadrants S/W and W/S combine the benchmark companies which underperform or outperform the target company on one of both indices: This then leads to the following benchmark portfolio mapping tool, which is, of course (by the very nature of the type of data used in the present analysis), limited to the technology characteristics of the companies (or units of analysis) examined:

Positioning ‘XXX’ viz. EPO–benchmarks	**Companies that are at a disadvantage in the overlapping part of their technology portfolio**	**Companies that have an advantage in the overlapping part of their technology portfolio**
Companies that have an advantage in their technology portfolio	**W/S**	**S/S**
Companies that are at a disadvantage in their technology portfolio	**W/W**	**S/W**

5. EMPIRICAL APPLICATION OF PATENT–BASED PORTFOLIO MAPPING

We have applied this methodology for S&T portfolio benchmarking to the top 20 companies in Belgium and Flanders in terms of patent strength. As mentioned, these companies account for more than 60% of the total EPO patent population in Flanders for the period 1978 until 1996. The companies were benchmarked against their respective groups of relevant, highly active peers in the EPO patent database, taking into account the necessary caveats when computing specialization indices as mentioned

Based on our computations we can map the relative position of the S&T portfolio of each of the companies against the benchmark group of controls. In Figure 26.3 the results for two of the major Flemish companies, Agfa NV (a photochemical company) and Janssen Pharmaceutica NV (a pharmaceutical company belonging to the Johnson & Johnson Group) are shown. The maps show a strong relative position of both companies, viz., their respective control groups, both in terms of total technology portfolios and in terms of the technology portfolio overlaps. Of course, these analyses can be further refined to examine and to cover specific subsets of the benchmark group in order to refine the relative position analysis. This step is easy to do since it only requires zooming in on specific areas of the S&T portfolio map.



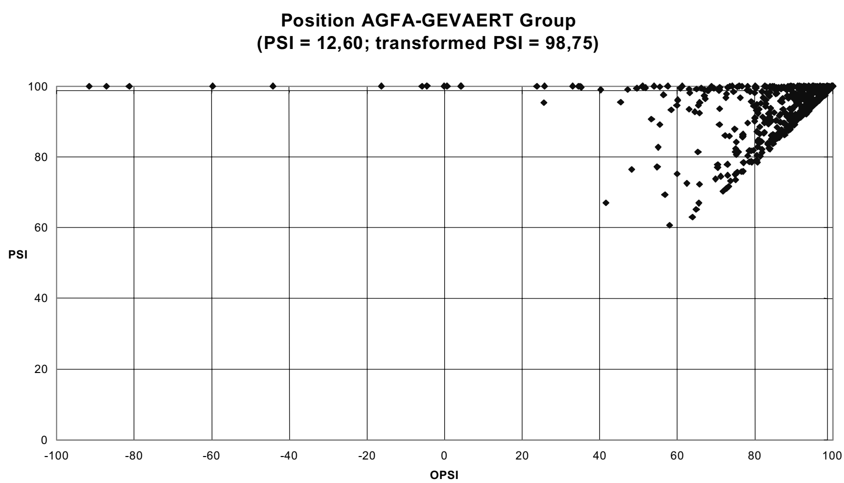


Figure 26.3. S&T Portfolio Benchmarks for 2 major Target Companies in the Flemish Region

6. CONCLUSION

In this chapter we have demonstrated the use of patent data to monitor an inter-organizational comparison of science and technology (S&T) portfolios taking a relative technology specialization perspective. It is obvious that other dimensions of portfolio analysis (e.g., relative market positions) can and could have been chosen as well.

In this chapter, however, we intended to focus on the use of patent data to map and to monitor relative technological positions at a relevant level of analysis (e.g., the most technology active companies in a regional innovation system).

As S&T portfolios have become instrumental in examining and monitoring the vitality of both institutions and regions in the innovation game that underpins their economic growth and development, the further development of benchmark tools and instruments should receive ample attention. As argued, those portfolios have to be monitored not only at the intra-organisational level, but also at the inter-organizational levels that are judged relevant for understanding relevant aspects and dimensions of the performance of specific systems of innovation.

Therefore the development of appropriate, easily used and transparent, benchmark indicators to assess the strengths and weaknesses of inter-organizational S&T portfolios is desirable. This has been the objective of the computational mapping described and developed in this chapter. This mapping approach is applicable to comparative analyses of organizational patent portfolios, as long as the organizations studied and the control group used are sufficiently patent active so that the Balassa Index can be applied.

REFERENCES

Abernathy, W.J., Clark, K.B. (1985). Innovation: mapping the winds of creative destruction. *Research Policy*, 14, 3-22.

Angelis, D.I. (2000). Capturing the option value of R&D. *Research Technology Management*, 43, No. 4: 31–34.

Antonelli, C. (1995). *The economics of localized technological change and industrial dynamics*. Dordrecht: Kluwer Academic Publishers.

Balassa (1961). An empirical demonstration of classical comparative cost theory. *Review of Economics and Statistics*, 4.

Boer, F.P. (2000). Valuation of technology using real options. *Research Technology Management*, 43 (4), 26–30.

Bone, S., Saxon, T. (2000). Developing effective technology strategies. *Research Technology Management*, 43, (4), 50–58.

Brown, S.L., Eisenhardt, K.M. (1995). Product development: past research, present findings and future directions. *Academy of Management Review*, 20, 343–378.

Cooper, R.G., Edgett, S.J., Kleinschmidt, E.J. (1997a). Portfolio management in new product development: lessons from the leaders – part I. *Research Technology Management*, 40 (5), 16–28.

Cooper, R.G., Edgett, S.J., Kleinschmidt, E.J. (1997b). Portfolio management in new product development: lessons from the leaders – part II. *Research Technology Management*, 40 (6), 43–52.

Cooper, R.G., Edgett, S.J., Kleinschmidt, E.J. (2000). New problems, new solutions: making portfolio management more effective. *Research Technology Management*, 43 (2), 18–33.

Cooper, R.G. (2001). *Winning at New Products.* Perseus Publishing.

Debackere, K., Luwel, M., Veugelers, R. (1999). Can technology lead to a competitive advantage? A case study of Flanders using European Patent data. *Scientometrics*, 44 (3), 379–400.

Dosi, G., Freeman, C., Soete, L. (1989). *The economics of technological change*. London: Pinter Publishers.

Floyd, C. (1997). *Managing technology for corporate success*. Gower Publishing Limited.

Foster, R. (1986). *Innovation: The Attacker's Advantage.* New York: Summit Books.

Girifalco, L.A. (1991). *Dynamics of technological change.* New York: Van Nostrand Reinhold.

Griliches, Z. (Ed.). (1984). *R&D, patents and productivity.* Chicago: University of Chicago Press.

Griliches, Z. (1990). Patent statistics as economic indicators: a survey. *Journal of Economic Literature,* 28, 1661–1707.

Jägle, A. (1999). Shareholder value, real options, and innovation in technology–intensive companies. *R&D Management,* 29 (3), 271–287.

Luwel, M., Noyons, E.C., Moed, H. F. (1999). Bibliometric assessment of research performance in Flanders: policy background and implications. *R&D Management,* 29 (2), 133–141.

Martino, J.P. (1983). *Technological forecasting for decision making.* North Holland.

McGrath, R.G., MacMillan, I.C. (2000). Assessing technology projects using real options reasoning. *Research Technology Management,* 43 (4), 35–49.

Meyer, M.H., Lehnerd, A.P. (1997). *The power of product platform.* New York: The Free Press.

Narin, F., Noma, E., Perry, R. (1987). Patents as indicators of corporate technological strength. *Research Policy,* 16, 143–155.

Narin, F., Rozek, R.P. (1988). Bibliometric analysis of US pharmaceutical industry research performance. *Research Policy,* 17, 139–154.

Narin, F. (1997). *Linkage between patents and papers: an interim EPO/US comparison*. Proceedings of the Sixth Conference of the International Society for Scientometrics and Informetrics. Jerusalem, Israel.

Nauwelaers, C. (2000). Towards a more interactive approach. *Innovation & technology transfer.* EC Publications, September: 12.

Nonaka, I., Takeuchi, H. (1995). *The Knowledge Creating Company.* Oxford: Oxford University Press.

Porter, A.L., Roper, A.T., Mason, T.W., Rossini, F.A., Banks, J. (1991). *Forecasting and technology management.* New York: John Wiley & Sons.

Perlitz, M., Peske, T., Schrank, R. (1999). Real options valuation: the new frontier in R&D project evaluation? *R&D Management,* 29 (3), 255–269.

Roussel, P., Saad, K., Erickson, T. (1991). *Third generation R&D – managing the link to corporate strategy.* Boston: Harvard Business School Press.

Schmoch, U., Hinze, S., Kirsch, N. (1992). *Analysis of the dynamic relationship between technical and economic performances in information and telecommunications sectors.* Fraunhofer Institute, ISI–Report Series.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *The Rand Journal of Economics,* 21 (1).

Vlaams Indicatorenboek Wetenschap, Technologie en Innovatie (2003). Steunpunt O&O Statistieken, K.U.Leuven.

Wheelwright, S.C., Clark, K.B. (1992). *Revolutionizing Product Development*. New York: The Free Press.

Chapter 27

PATENT PROFILING FOR COMPETITIVE ADVANTAGE

Deducing Who Is Doing What, Where, and When

Alan L. Porter[1] and Nils C. Newman[2]

[1]*Search Technology, Inc., and Georgia Tech , GA, USA. E-mail: aporter@searchtech.com*
[2]*Search Technology, Inc. and IISC, USA*

Abstract: This chapter introduces text mining of patents in support of technology management. Technological innovation models point to empirical measures that relate to prospects for successful commercialisation. We present an 8-step process for analysing entire patent sets on a given topic to generate such 'innovation indicators.' We illustrate for the case of fuel cells.

1. PATENTS, TECHNOLOGICAL CHANGE, AND TECHNOLOGY MANAGEMENT

A patent constitutes a milestone in the advance of a given technology. However, it is only one piece in a larger puzzle of technological innovation. A patent reflects 'invention' — the birth of a new idea — but 'innovation' entails combining ideas with a suitable business strategy combined with other factors to achieve commercial success (Schumpeter, 1975). The lag between invention and innovation can be sizeable. Thus patents provide an astute observer with a window on the future. If you know what you are looking for you can deduce the technological trajectory of your opponents and predict their strategic moves before they make them. This game of ascertaining 'who is doing what, when' is Competitive Technological Intelligence (CTI). Patent analysis plays an important role in CTI by helping to discern patent activity patterns with implications for innovation.

Patent analysis takes many forms, with important distinctions between micro- and macro-analyses (Trippe, 2003). In the private sector, for instance, intellectual asset management groups probe deeply to understand individual

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 587-612.

patents and their associated claims. In the public sector, policy studies may analyse an entire nation's patenting combined with R&D publication activity to understand better its science and technology ('S&T') capabilities.

This paper targets a special intermediate level of analysis. The unit of analysis is not individual patents, but collections of patents on a specific topic. This 'patent profiling' seeks to find patterns in the development of particular technologies. The resulting CTI can help characterise current development efforts and also help anticipate future developments, thus generating practical, timely information in support of specific technology management decisions[1]. The term 'Patent Analysis' is too broad and unspecified to describe this, while 'Patenting Profiling' is slightly narrow. To better convey our intent to exploit science & technology information on behalf of technology management, we use the term 'Tech Mining' throughout this paper to signify this intermediate level of analysis (Porter and Cunningham, in preparation).

What one emphasises in Tech Mining (**'TM'**) depends strongly on the intended uses of the intelligence to be generated. If one stands back and considers the approaches presented in this Handbook, some mainly serve scholarly inquiry; others, public policy interests; and still others, company technology management. This paper focuses on the last of these seeking ways to generate practical, timely information in support of specific technology management decisions. As such, TM of patents can support decisions concerning:

- R&D program management;
- Mergers and acquisitions;
- New product development;
- Intellectual asset management;
- Management of technical human resources;
- Technology Foresight;
- Strategic planning.

It is important to note the word 'support.' Patents, of course, do not tell the whole story. An astute observer does not rely on only one observation. Analysis of R&D publication activity and business information (e.g., articles, product announcements, market assessments) enrich TM. Moreover, all these empirical analyses must be complemented by expert opinion.

The target of TM patent information for CTI is to generate actionable knowledge that can support decision making. The prime end users are

[1] A corresponding approach using R&D publications abstracts, 'research profiling,' can provide a bird's eye view of scientific research domains (Porter et al., 2002).

technology managers and professionals who must determine suitable courses of action. In the past, such decisions have relied overwhelmingly on tacit knowledge — usually in the guise of 'managerial intuition.' With the increased availability of rich information resources, we can do better. This paper suggests ways of accomplishing this objective through TM of patent abstracts.

1.1 Conceptual Foundations — Innovation Indicators: Designing Actionable Knowledge

When we speak of TM of patents we are not talking about search and retrieval. The goal is not to discover an individual patent. We intend to capture the complex patterns hidden within a collection of patents (Teichert and Mittermayer, 2002). These patterns relate the patents within the target set to stages within the technology innovation process.

The notion that technological innovation processes have a reasonable degree of regularity is essential to TM. Technology intelligence presumes a degree of orderliness to the innovation process. Emergence of new or improved technologies depends on successful completion of the innovation process — "any system of organised activities that transforms a technology from an idea to commercialisation"(Souder, 1987). We draw upon various innovation and technological change models to generate a set of concepts to help gauge the prospects of particular technologies becoming successes — i.e., 'innovation forecasting' (Porter and Watts, 1997).

Successful innovation relies on internal organisational factors as well as external factors — technology characteristics, market forces, and socio-economic-political climate.

Extensive research documents the factors that either promote or inhibit successful product, process, or service development, commercialisation, and spread. Many researchers have performed post mortem assessments of technology transfer activities, technology diffusion, and technology substitution processes to characterise significant factors and recommend managerial practices which promote success in new product technology innovation [c.f., Millson et al., 1992; Mahajan and Muller, 1996). We have scavenged 'innovation success' concepts from various sources. In particular:

- Michael Porter's (1985) 4-factors framework highlighting the importance of various competitive forces;
- Souder's (1987) identification of organisational factors relating to technical or commercial success;
- Dunphy et al.'s (1996) juxtaposition of the factors of an innovation funnel;

- Smith's (1992) specification of levels and forms of substitution;
- Modis' (1993) observations on compatibility with infrastructure and complementary products;
- Anderson and Tushman's (1990) evidence on the interplay of industry participants;
- Dror's (1989) use of patent information to infer design standards.
- Metcalfe's (1981) technology diffusion considerations;
- Cohen et al. (1979) and Rogers (1983), each identifying sets of factors contributing to technology transfer;
- Souder et al. (1990) on roles of sponsor and adopter.

We draw on these concepts to develop 'innovation indicators' — empirical measures based on mining S&T information that can contribute to technology management (Porter and Watts, 1997). We divide these into three main families:

- *Technology Maturity* — i.e., stage in the life cycle — advance along the development pathway (possibly an 'S-shaped' growth curve), growth rate, and the development of critical component technologies and system technologies into which the innovation fits.
- *Innovation Context* — i.e., the technical, economic, environmental, and social context influences on the technology in question, including: alternative technologies to achieve similar functions, pertinent regulations & standards, and prevailing institutional interests.
- *Market Prospects* — i.e., product value chain — considering primary and secondary applications, and sectoral & geographical dispersion.

Analysis of patent information contributes most to maturation indicators. Complementary TM on other information sources more fully taps into contextual and market indicators.

Given this rich set of targets, the secret for success in TM is to determine what information sources and, specifically, what types of patent sources give you data from which you can create these indicators. Toward this end Ernst (2003) presents a rich array of possible indicators aimed at actionable technological intelligence, including:

a) Technological Emphasis: Compare an organisation's patenting activity (applications or patents granted) among 'fields' (e.g., use International Patent Classes or your own markers of domains of special interest).
b) Technology Share: For a given field, compare various organisations' extent of patent activity.
c) Rate of Technological Growth: Compare recent versus earlier activity levels.

d) Patent Quality: Adjust the amount of patenting to take into account the quality.

We will later demonstrate how data can be exploited to estimate these indicators and then draw broader implications about firm behavior (see also Granstrand, 1999).

Contrast this 'TM' perspective with a 'data mining' one. Data miners start with an information resource and bring to bear analytical tools (e.g., basic tabulations, statistics, applied artificial intelligence) to glean as much information as possible. We begin at the other end — focus on what technology managers need to know, reflected through our conceptualisation of factors influencing innovation, back to exploit the available information resources.

1.2 Multiple Information Resources and Analytical Approaches

Constructing innovation indicators as part of TM can, and should, draw upon multiple information resources. When working on public sector projects, our innovation indicators development draws heavily upon research publication and conference paper abstract databases, such as *Web of Science, MEDLINE, INSPEC,* and *EI Compendex* (c.f., "technology opportunities analysis" at http://www.tpac.gatech.edu). We also use more cluttered sources such as the *Business Index*. In addition, the Internet offers a rich, but messy, repository of S&T information. The databases offer some quality assurance and focus (i.e., one knows the nature of the sources), but the Internet offers currency (vs. the time lags inherent in compiling information into databases from sluggish publication and patenting processes). However, when addressing many private sector technology management issues, patents are often the only reliable public source of information about competitor innovation activities.

Within patents there are many different sources for retrieving records on a technology of interest. Patent aggregators such as Delphion and MicroPatent provide coverage of multiple patent authorities. Individual patent offices such as the U.S. Patent and Trademark Office and *European Patent Office* offer their own search systems but use different classification systems. World Patent Index ('DWPI' for short) offers value–added patent information by rewriting the patent abstracts into comprehendible language. They also classify records within their own classification and record structure.

For some purposes citations (referencing by patents to prior patents and to non-patent sources) can be very valuable. The basic nature of citation analysis is to identify those patents receiving significant numbers of citations

(Trajtentberg, 1990; Mogee, 2003). For a target technology, patent citation analyses can distinguish influential patents, prominent organisations, and/or leading inventors. One can gauge the pace of innovation by measuring median time lags from citing applications to the grant dates of cited patents. Tabulating the percentage of citations by an organisation to itself can indicate whether it is a pioneer and whether it is assertively protecting its intellectual property ('IP') in the target arena. Mapping 'citation trees' over one or more generations (http://www.mogee.com/) can elucidate intellectual parentage. As with most S&T data, distributions tend to be extremely skewed with a few patents accruing large numbers of citations and many patents receiving none or almost none. Citation analysis can focus on individual patents or on clusters.

Chen (2003) shows patent 'landscape' maps created by Boyack et al. (2003) changing over time. These aid CTI in tracking competitor interest evolution. Patent citation of scientific publications is increasing dramatically (Lane and Makri, 2000). Patterns in these citations bear on S&T policy deliberations (Narin et al., 1997).

The types of indicators and the types of data sources vary greatly depending on the technology management questions being asked. The applicability of patent data and the TM process will also vary from technology to technology depending on the patenting behavior of specific industries. For example, software developers tend not to patent, while many chemical compounds are detailed in streams of patents.

2. THE NITTY GRITTY OF TECH MINING ('TM')

As noted previously, TM should start with a well-defined issue being addressed on behalf of well-defined target users. Put another way, value derives from successful application to help resolve managerial issues, not from scholarly insight gained. This emphasis on *utility* balances traditional scholarly concerns with *validity*. We address both.

Validity first requires that suitable data be collected to ascertain activity patterns for the technology under scrutiny. Information search must reflect proper scope. The typical targets of information retrieval are recall and precision. TM wants to capture a high percentage of relevant records (i.e., high recall), but missing some is not devastating, as it could be in legal claims or patentability investigations. Similarly, one wants as much 'signal' and as little 'noise' as possible (i.e., good precision). But again, TM is not too sensitive to the presence of noise. During the course of the analysis any sizeable body of irrelevant records will be uncovered and can be set aside.

TM data sensitivities thus differ from those in other information retrieval exercises.

TM analytical objectives also differ from those of other inquiries. They include:

- 'Domain profiling' — i.e., providing an overview of R&D efforts relating to the target technology in their entirety (Porter et al., 2002).
- Depiction of "who is doing what" — i.e., R&D and intellectual property protection activity patterns, spotlighting leaders in particular facets, etc.
- Capability of 'zooming in' — e.g., to follow up on an interesting lead by quickly getting to details such as the patent records generated by a putative knowledge network.
- 'Knowledge discovery' — i.e., making links among concepts, technologies, or methods of which you were unaware.

To generalise, these objectives are not highly sensitive to noisy data points. Empirical analyses can discern a surprising amount about a technology, even though the analysts know little about the area. However, there is also danger of misunderstanding. Patent information presents only part of the picture and even that information may not accurately reflect R&D results or innovation purposes. The best defence is to engage knowledgeable persons in reviewing interim TM analyses. Experts can help properly scope data searches, flag data that ought to be excluded, and catch gaps that require search augmentation.

Much of the indicator development in TM entails comparison. Within a technological domain, we want to know the more active topics, prolific inventors, and organisations assigned the most patents. Across domains, we compare relative activity patterns, such as patenting trends. Benchmarking one firm against others is often illuminating. Such comparisons are not usually cast in terms of statistical significance, but it behoves analysts to strive for unbiased comparisons (c.f., Campbell and Stanley, 1963) and to present results fairly. For instance, rather than say "company A has twice the number of patents of company B" when the numbers are small, present the actual numbers (e.g., 2 vs. 1).

It is also vital to inform users of how to interpret TM findings. In one case, a colleague at the U.S. National Institutes of Health presented a domain topic map to an audience of researchers. One objected to the placement of her work, thereby destroying the credibility of the entire analysis. In fact, the mapping had been done correctly and the placement could be explained. To avoid such devastating situations, be sure that users understand the broad-brush nature of research profiling and knowledge discovery in such analyses. TM generates insights and new perspectives rather than absolutes.

Utility issues will be revisited later when discussing applications. In terms of setting forth TM premises, the key is to include utilisation from the onset as a major objective. That should prompt the analysts to identify target users, ascertain their information needs, and strive to engage them in scoping and analyses. Doing so will greatly bolster the prospects of success.

At this point, let us pursue a case analysis to illustrate these TM ideas with a specific set of patent analyses.

3. FUEL CELL PATENT ANALYSIS

TM patent analysis consists of eight steps (which can be compressed or expanded as appropriate):

1. Issue specification and data source(s) selection;
2. patent search and retrieval;
3. data cleaning;
4. basic analyses;
5. advanced analyses;
6. information representation;
7. interpretation;
8. utilisation.

We briefly illustrate each.

3.1 Issue Specification and Data Source(s) Selection

For the purposes of demonstration, we selected *fuel cells* as the sample topic of an emerging technology with broad promise and a considerable patent record base. A fuel cell converts hydrogen and oxygen into water, producing electricity and heat in the process. It is thus an electrochemical device, somewhat like a battery, but it is 'recharged' with hydrogen and oxygen instead of electricity. Unlike most actual TM, the present purpose is simply to illustrate a sample analysis. With this simple statement, we have bypassed one of the trickiest steps in the TM process — getting management to specify the problem. We will discuss techniques to enhance managerial buy-in to TM latter in the paper.

Data for this analysis is from a search in *Derwent* (DWPI) via *Dialog* (a leading gateway to over 400 databases)[2]. We draw occasional comparisons to analyses of 11,764 fuel cell publication abstract records from 1967 to

[2] We thank Thomson Derwent and Thomson Dialog for providing the data.

2002 retrieved from *Science Citation Index* and *INSPEC* (Porter and Cunningham, in preparation).

3.2 Patent Search and Retrieval

A search on 'fuel cells' in DWPI was conducted in March, 2003. The 'fuel cells' search used:

– International Patent Classification (IPC) code H1M-8 (fuel cells) (*or*);
– The phrase 'fuel cell(s)' in patent family title or abstract.

The search was not restricted to a date range. The earliest record is from 1967.

The search yielded 23,836 records (before we eliminated patent families that consisted only of Japanese patents). Had we just searched on the H1M-8 classification code we would have missed 5,245 of these (22%). Had we just searched using the terms, we would have missed 7,920 (33%). More so than in publication abstract database searching, patent searching benefits from a combination of index–based search (using class codes) with Boolean term searching. Classifications are not globally standardised, although efforts are underway to harmonise Japanese (JPO), European (EPO), and United States (USPTO) classifications. Also, note that national patent offices do not assign International Patent Classification (IPC) codes uniformly. Codes are also updated frequently.

A more thorough search might have added synonyms or related terms for 'fuel cells.' In practice, it is often helpful to iterate initial searches to sharpen the focus by scanning the retrieved records for related terms to augment the term–based search. One may also identify extraneous terms whose exclusion could reduce noise significantly. However, as noted, TM is not unduly sensitive to noisy data.

3.3 Data Cleaning

Data manipulations and analyses used Search Technology's *VantagePoint* software ('VP' for short; http://www.theVantagePoint.com). Closely related versions of the software are available as *TechOASIS* for U.S. Government use and *Derwent Analytics*, tailored to expedite analysis of DWPI.

Many of the DWPI records reflect patents filed only in Japan. Japanese patents always present an issue as a result of two policies of the JPO: the JPO has a two tier examination system in which full examination does not occur until the patent is challenged and the JPO accepts multiple patents with single claims when other systems would have the inventor consolidate the patents into one patent with multiple claims. To mitigate these issues, we chose to reduce the fuel cell set to 9,724 records — only including Japanese patent families that also contain patents in a country other than Japan. In other words, the Japanese patent has been vetted by another country's patent system. The 9,724 patent family abstracts reflect 31,559 patents (or 3.3 patents per family). However we see 34,073 instances because some patents appear in multiple families (closely related patents).

Data cleaning can dramatically improve the quality of analyses. Software such as VP enables cleaning by removing duplicates and combining closely related entities (e.g., variants on an author's or inventor's name). VP provides standard thesauri, along with capabilities of 'growing' one's own specialised thesauri (continually augmenting over time), and 'fuzzy logic' modules to help match terms. Scripting the basic cleaning steps makes these easy to apply, although human checking may be warranted.

3.4 Basic Analyses

We are now ready to exploit the patent abstract records. Following van Raan (1988; also http://www.cwts.nl/), let us distinguish this section's first order (lists) and second order (matrices) analyses from the next section's third order (clustering and other pattern discerning processes).

Making lists that tally content of particular fields over a large set of records is simple, but it is also valuable in TM. Lists are easy to understand, and that facilitates use in decision support. Matrices simply combine two lists of interest. Much of the resulting information can be summed up as indicating "who is doing what." This section selects a few lists and matrices to illustrate TM uses.

Table 27.1 lists some of the 45 fields, deriving from the DWPI records, available in our VP dataset. Derwent has changed its specialisations (e.g., Abstract ADVANTAGE) over time so such fields must be used cautiously. The fields present various types of information: content (abstracts, titles), classification (Derwent Classifications, File Segment, and others not shown, such as IPC code), patent inter-relationships (family), year, and country.

Table 27.1. Sample Patent Abstract Fields

Field	*Number of Items*
Raw Record	9,724
Abstract	9,437
Abstract ADVANTAGE	5,948
Abstract Phrases (NLP)	118,683
Basic Patent Year	39
Derwent Classifications	278
Family Member Countries	42
Family Member Years	39
Family Member Years (most recent)	37
Inventors	10,112
Patent Assignees	3,311
Priority Countries	41
Priority Years (earliest)	44
Tech Focus	1,892
Title	9,631

The data present three patent typologies: ‘Family’ (a group of related patents), ‘basic’ (the original patent entered into DWPI to which the others are the same (but filed in another country) or closely related), and ‘priority’ (the first patent filed). Here we have only imported country and year information for these types, but one would determine what information is useful for particular purposes. For instance, CTI can gain from plotting the distribution over time of a company’s ‘priority’ vs. ‘family’ patents (overall or on a specific topic). The lag between first patenting applications and patents issued can tell about that company’s patenting skill, global market interests, and strategies (e.g., ‘submerging’ patents by ongoing modification of an application to delay patent issue).

Suppose we are examining Western European technological strength in fuel cells. To illustrate use of a data subset, we select 1) automotive–oriented fuel cell patents, 2) dating 2000–2003, for 3) assignees in Western European countries — yielding 278 records. Table 27.2 gives the ‘Top 10’ companies. Note that many are automotive companies, but others, such as Siemens, have related fuel cell patents in this domain as well as in other areas.

Table 27.3 shows the family member distribution for basic patents from the top 10 European automotive fuel cell companies. This table indicates the patent protection behaviour for European firms. For instance, 47 of Xcellsis’ 49 patents show German priority. Xcellsis has then gone on to file with EPO on 35 of these, and for U.S. protection on 34, but only 11 in Japan. We shall explore Xcellsis further shortly.

Table 27.2. Leading Automotive–oriented, Fuel Cell Patent Assignees in Europe, 2000–03

Patent Assignee	*Records*
XCELLSIS	49
DAIMLERCHRYSLER	40
Siemens-Westinghouse	22
MANNESMANN	20
VOLKSWAGEN	18
EMITEC	17
RENAULT	16
DBB FUEL CELL ENGINES	15
VALEO KLIMASYSTEME	7
BOSCH	6

Table 27.3. Distribution of Priority Country by Family Member Countries for European Assignees

		Family Member Countries									
	Basic Patent Assignee	DE	EP	WO	US	JP	FR	AU	CA	GB	
49	XCELLSIS	47	35	8	34	11	1	1	2		CN
40	DAIMLERCHRYSLER	40	12	4	13	5	3			1	
22	Siemens-Westinghouse	19	10	20	8	5		2			
20	MANNESMANN	20	5	14	1	2		3			2
18	VOLKSWAGEN	18	2								
17	EMITEC	17	4	17	4			6			
16	RENAULT		5	3		1	16	1			
15	DBB FUEL CELL ENGINES	14	9	2	11	11		1	1		
7	VALEO KLIMASYSTEME	6	3		3	4	3				
6	BOSCH	6	1	3	2					2	

Table 27.4 profiles patent activity for the five companies that published the most on fuel cells (based on analysis of the R&D publication abstracts data). There are several points to notice:

– Siemens and Westinghouse each show extensive patenting; for some purposes, one might analyse these merged companies separately.
– The leading publishers are not the same as the leading patenters — Tokyo Electric shows minimal patenting, whereas International Fuel Cells shows 251 patents in the data set and Honda 186 (not shown here).
– Siemens-Westinghouse shows extensive recent patenting, unlike some of the other companies.
– Distinctions in emphases do not appear generally from the Derwent patent classifications, but one could explore activity in particular classes, such as X21-Electric Vehicles.
– Certain inventors are quite prominent (not shown) — one might well explore ‘knowledge networks’ in these companies.

Managers report such profiles customised to address key aspects of the leading players or topics very informative.

Table 27 4. Patent Activity of the 5 Leading Research Publishing Companies

Assignees	*Family Years (most recent)*	*Priority*	*Derwent Classes*
Siemens-Westinghouse	2002 [137]; 2001 [28]; 2000 [26]	Germany [374] USA [177] EPO [23] WIPO [6] Japan [2]	X16-Electrochemical Storage [529]; L03-Electro-(in)organic [347]; A85-Electrical applications [50]
UTC	1974 [17]; 1989 [16]; 1988 [15]	USA [251] France [1] WIPO [1] Japan [1] EPO [1]	X16-Electrochemical Storage [208]; L03-Electro-(in)organic [200]; A85-Electrical applications [64]
Hitachi	1987 [9]; 1992 [7]; 1993 [7]	Japan [78] WIPO [2] UK [1] Germany [1] Netherlands [1]	X16-Electrochemical Storage [72]; L03-Electro-(in)organic [55]; A85-Electrical applications [14]
Mitsubishi Electric	2002 [8]; 1998 [7]; 1990 [7]	Japan [49] USA [14]	X16-Electrochemical Storage [59]; L03-Electro-(in)organic [41]; X21-Electric Vehicles [6]
Tokyo Electric	2001 [1]; 1994 [1]; 1996 [1]	Japan [3]	X12-Power Distribution [1]; X13-Switchgear [1]; X16-Electrochemical Storage [1]

3.5 Advanced Analyses

Various activity patterns and relationships can prove informative. This section introduces approaches for combining information in ways not plausible without text mining. For instance, we could examine 'where' a company patents, and how this changes over time, to gauge its market and intellectual property emphases. Several of this section's examples include Ballard Power Systems patenting. Ballard has 40 or more patents in seven patent offices. Their Japanese patenting peaked in 1999, dropping to a single patent in 2001–2002. In contrast, they have 36 EPO and 52 U.S. patents in these recent two years; this may speak to their relative market intents.

Temporal information is vital to track technological maturation. Suppose we want to check companies' general fuel cell patenting activity *trends*. Figure 27.1 suggests how extraordinarily active fuel cells have become by comparing companies using most recent family patent years. These include the top R&D publishing companies (Table 27.4), except for Tokyo Electric which does not patent much. Siemens and United Technologies (UTC) are leaders in both patenting and publishing. We add Ballard, International Fuel Cells (IFC) and Honda who are leaders in patenting, but do not publish as actively. UTC and Siemens-Westinghouse were very active in prior years (not shown for space reasons here). Recently, UTC seems to have largely ceased this effort; as with much TM, this raises questions for additional 'detective work.' Honda, in contrast, has escalated its fuel cell patenting remarkably the last few years — another factor of note for further CTI exploration.

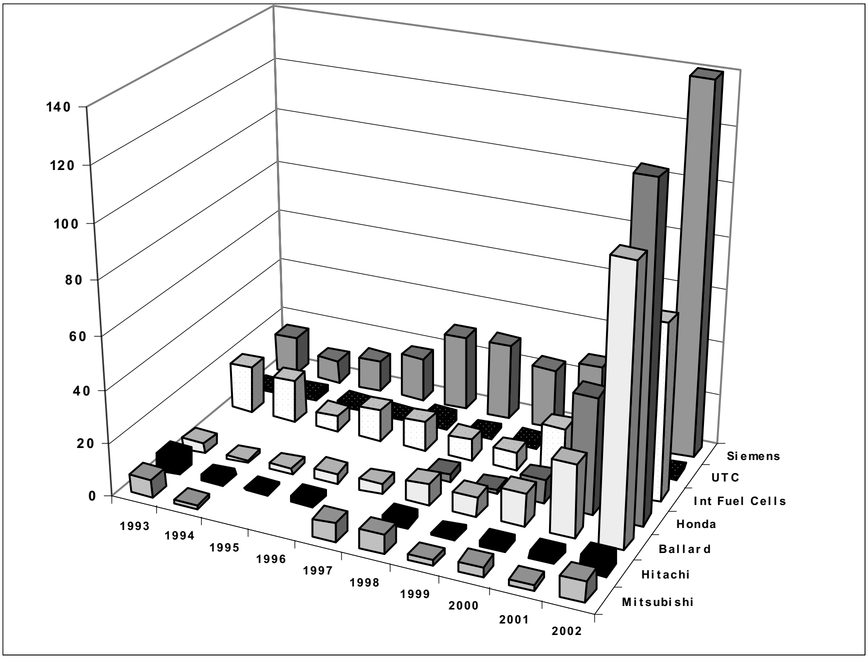


Figure 27.1. Fuel Cell Patenting Trends for Selected Companies

Other trend analyses can contribute to technology intelligence and forecasting. For instance, we might compare patenting for each of the five main fuel cell types. We could project trends into the near future, overall or for particular technologies or companies, by fitting growth curves to these time series data. Grouping data into 'time slices' can display significant shifts. This can work well with sequences of cluster maps (introduced shortly). Another way of combining topic and trend information is by preparing 3-D surface maps. These can be constructed for a sequence of time slices, marking one organisation's publications or patents to see the progression of its interests by topic, over time (c.f., Boyack, 2003; Borner et al., 2003; Porter and Cunningham, in preparation).

One can spot new or hot topics by *list comparison* of terms for the most recent time period with those used in prior times. Expert review of such lists helps identify candidate important topics. To illustrate, Ballard Power Systems is a leading fuel cell company. Suppose we want to ascertain its current thrusts. Using VP, we create one file of Ballard's 49 2002 patents and another of its 73 patents from 1999–2001. VP's natural language processing (NLP) provides lists of title noun phrases in each file. Using a list comparison function, we then browse the title phrases new in 2002 to obtain clues about possible Ballard initiatives. The new title phrases include, for instance: annular catalytic reactor tube, catalyst nano-dispersed, and gas purification device.

Mapping related terms is, perhaps, the most compelling way to identify patterns. Co-occurrence of entities across records provides the basic information from which to deduce possible relationship. Widely used versions include: co-citation, co-authoring, and co-word. Co-citation builds putative linkages based on papers or patents citing others together. Henry Small and others have applied co-citation to identify and track research fronts. (c.f., Small and Griffith, 1974).

Co-authoring (or co-inventing) provides sociometric information about individuals collaborating. Figure 27.2 shows a co-inventor map for Ballard's 176 fuel cell patents since 1993. In this VP generated map, node size reflects number of patents. Location provides a crude sense of possible relationship using multidimensional scaling; the axes have no inherent meaning. The connecting lines indicate relative strength of association using a path–erasing algorithm (Zhu and Porter, 2002). Listing Ballard's 274 inventors finds Wilkinson with 47 patents and another 11 with 8 or more (an arbitrary cut-off for this illustration). A co-occurrence matrix of inventors by inventors gives us details about who works with whom, but the map (Figure 27.2) helps us perceive relationships among the top 12 inventors more readily:

- Wilkinson is central to most of the invention in this major fuel cell company; he has co-patented with 9 of their other 11 leading inventors;

– Voss, the second leading inventor, works heavily with Wilkinson, but seems to reflect a second node of activity; note that Barton collaborates with Voss, but not with Wilkinson.
– Stone (and Steck, with 7 patents, not shown) collaborate closely, and work with three others, who together form a distinct group — none of these five seem to work with any of the other leading Ballard inventors.

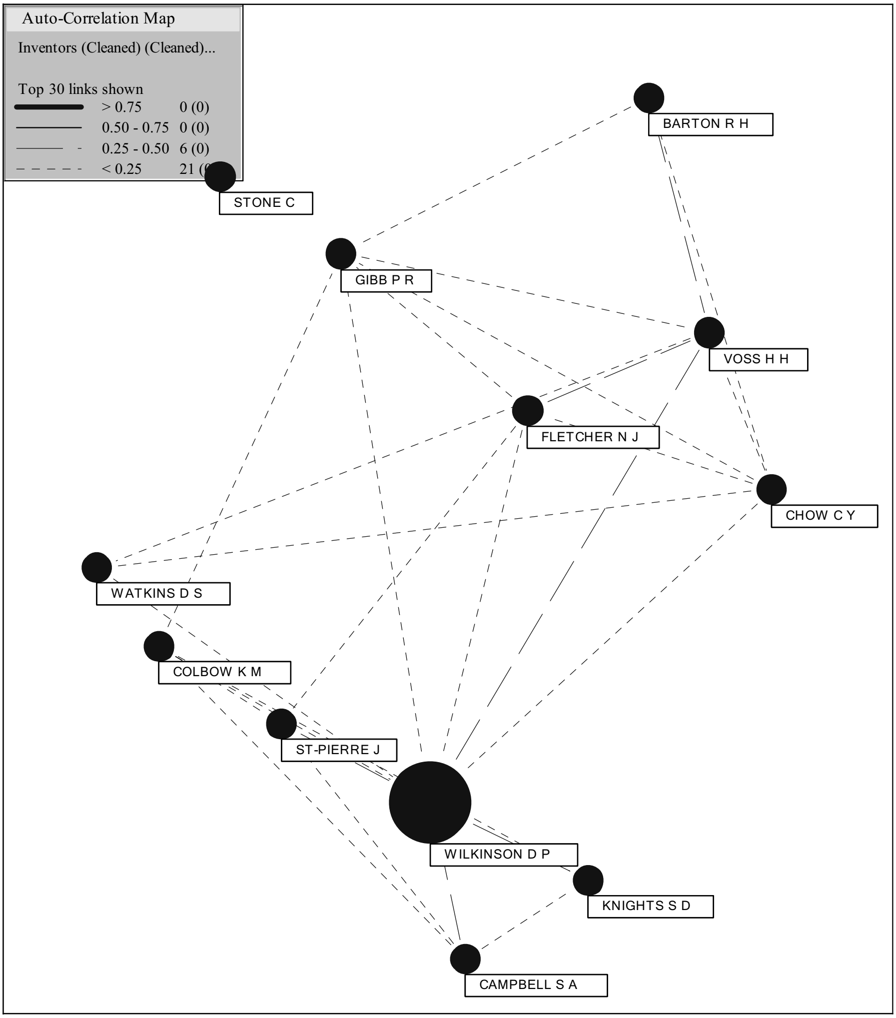


Figure 27.2. Knowledge Network within Ballard Power Systems

This ‘knowledge network’ analysis of Ballard’s patent activity could be pursued by examining the topics of the respective patents and the patent teams' recent emphases. This intelligence could help another company forecast Ballard technology developments and possibly their market intentions. It could set the stage for overtures to collaborate or even recruiting forays to entice away pivotal technical knowledge.

Similar analyses can be carried out to see which companies patent together. Such an analysis of the leading automotive–oriented fuel cell patent assignees in Western Europe (recall Table 27.3) identifies several links of possible interest (Figure 27.3). Suppose our company is trying to understand this competitive technological environment. We note that

– 3 of the companies active in European automotives are among the overall leaders in fuel cell patenting - Daimler-Chrysler, Siemens-Westinghouse, and XCELLSiS;
– Emitec and Mannesmann have strong ties to Siemens;
– DBB and XCELLSiS are linked to Ballard.

Such results cannot, of course, tell the whole story. For instance, collaborators may determine that only one obtains the patent while the partner receives a suitable license or other appropriate share in the development.

We follow up interesting pointers, such as this last one, through further analyses. By charting Assignees by Assignees, we see that Ballard and DBB joint inventions peak in 1998 (8 patents) and end in 1999. That prompts us to look at all 84 DBB patents and we see that these rise and fall abruptly. We then look at EXCELLiS' profile, noting that it is peculiarly complementary to DBB — showing a dramatic upsurge to 68 patent filings in 2000! A Google web search uncovers that Daimler and Ballard collaborated from 1993–97. In 1997 they expanded this collaboration with two jointly owned companies, XCELLSiS, formerly DBB Fuel Cell Engines, and Ballard Automotive. Furthermore, in 1997 Ballard and Daimler expanded their alliance to include Ford Motor Company, with a third venture, Ecostar, to develop electric drives for electric vehicles. So patent analysis, augmented with follow on web searching, yields valuable technological intelligence on "who is doing what with automotive fuel cells."

This vignette also hints at complexities in tracing company interests through patent analyses. Daimler's fuel cell patenting appears to drop sharply from its peak in 1999 (as priority year), but, as noted, they have initiated several joint ventures. This implies more, not less, commitment by them to develop automotive fuel cell applications.

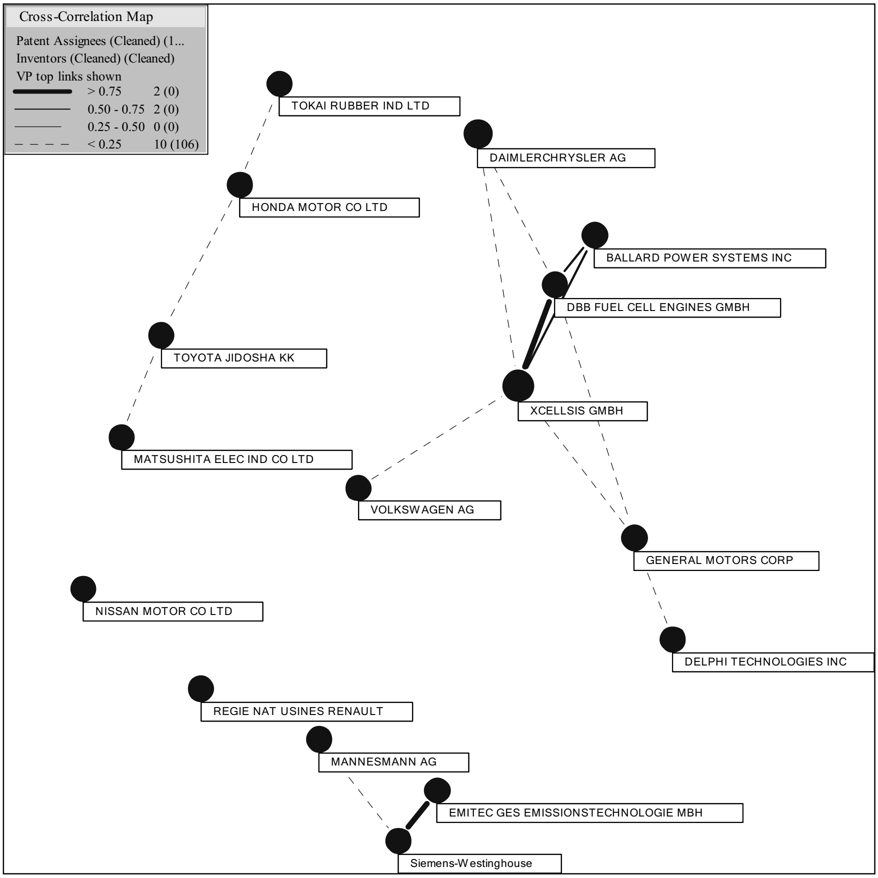


Figure 27.3. Fuel Cell Co-Patenting

Co-term analysis can also discern topical thrusts. We most often use keywords (subject index terms) for this. VP uses principal components analysis (PCA) to 'cluster' terms that tend to occur together (Zhu and Porter, 2002). VP can also group title and/or abstract phrases or classification codes. In a variation, 'principal components decomposition' applies an optimising routine to group records with shared content. This offers an inductive way to classify R&D, as opposed to an imposed (deductive, index) approach.

Figure 27.4 shows a PCA map of Derwent manual codes. Each cluster is named after its most central manual code. This tends to form patent application domains (e.g., nodes such as 'textile applications,' 'polymer applications', mixed with technologies.) Note the pull-down examples that display the 'high loading' codes that form the PCA factor. Cluster maps are best used as navigation tools, not static representations as presented here. In a dynamic environment, researchers can further explore relationships to investigate topics of concern for particular purposes. For instance, a little exploration is required to cover broad topic areas such as automotive applications that intersect a range of manual codes. Navigation within the clusters can also provide researchers with more detailed information within specific clusters. For example, a researcher could explore within the Textile Applications cluster to provide a detailed understanding of the content of the cluster.

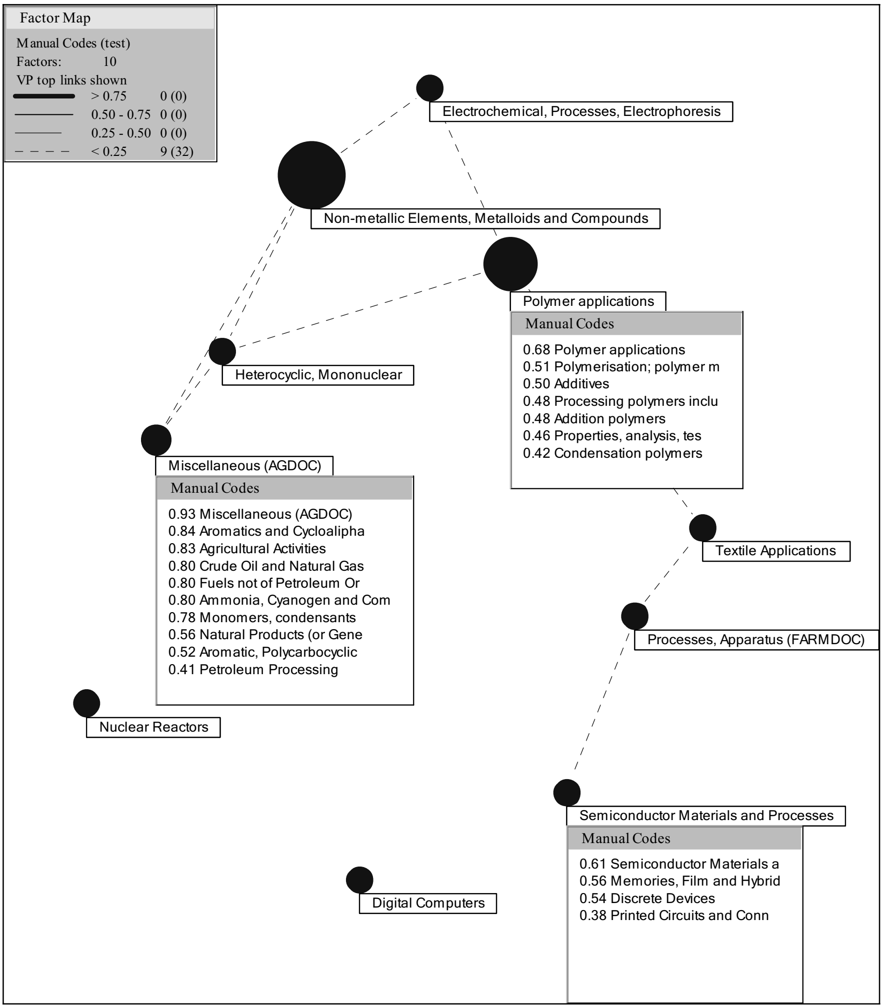


Figure 27.4. Fuel Cell Topical Emphases Based on Derwent Manual Codes

3.6 Information Representation: Building Indicators

How can one best convey TM analytical results? It depends, of course, on the training and interests of the intended users (e.g., technical specialists tend to want analytical detail while managers and officials tend to favour more integrated, holistic information. The patent analyst, further, needs to know individual user preferences, implying serious attention and effort to learn them. These should determine what information to provide and also how to present it — text vs. charts, tables vs. figures, more vs. less interpretation.

Innovation indicators (introduced in Section 1.1) lend themselves to graphical presentation. These consolidate analytical results (Sections 3.4 and 3.5) to address specific technology management concerns. Indicators are best considered from the end use back toward the analyses and data. For instance, imagine that the TM of fuel cells is being performed for Volkswagen to help determine whether to pursue development of a particular fuel cell technology for a particular automotive application. Different aspects of this issue involve various corporate functions. To give the flavour, consider possibilities for the three types of innovation indicators (Section 1.1):

– Technological Maturity — trends in development of critical components or systems pertinent to Volkswagen's target application.
– Innovation Context — identification of competitive technologies, assessment of standards, indications of institutional interests (pro and con), and prospect of regulations.
– Market Prospects — evidence about market existence for the target application (user needs assessment), market size and rate of change, applications noted, sectors engaged, and benchmarking Volkswagen's competitive position versus key players.

Indicators can be specialised further to address particular technology management needs. For instance, suppose the purpose of the Volkswagen patent analysis is to support a decision about whether to acquire a small company. The tech miners could prepare an integrated presentation drawing on empirical and expert information to make a recommendation. Patent findings might include: the intellectual property (IP) strength of the company (map of the patent portfolio's coverage), IP currency (maintenance of patent protection), IP coverage (locations), and knowledge network robustness (characterisation of the invention teams and whether core players remain with the company).

Such results can be elevated to another level for quick assimilation and comparison. We call these 'scorecards' — visualise simple, stacked horizontal bars, each scaled 0-100:

– Science Base: % of patents referencing publications;
– Technological Maturity: Current location of R&D along an S-shaped growth curve from onset to saturation;
– Competitive Entry: Slope of the trend in new companies initiating patenting in the most recent 1/3 of the years under study;
– Competitive Exit: % of top assignees who have left the domain;
– Diffusion: Change in number of new IPC codes over the most recent 1/3 of the years under scrutiny.

Such a scorecard can provide busy executives with a quick sense of activity in the domain under scrutiny (e.g., a sub-set of fuel cells).

Figure 27.5 shows another high-level representation. Imagine Volkswagen executives being briefed on their position vis–à–vis Daimler and Ford in automotive fuel cell technology. We combine two of Ernst's five metrics (Section 1.1), Technological Status (possibly measured by patent activity) and Rate of Change. The figure makes up values for three companies to suggest the value of such simple benchmarking.

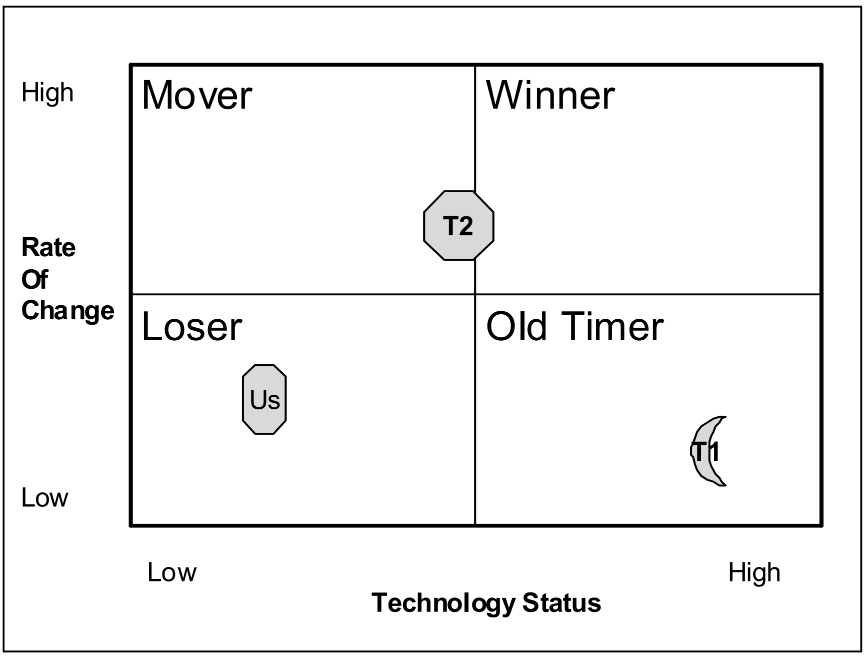


Figure 27.5. Competitor Profiling

3.7 Interpretation

We have noted the range of interpretive possibilities from minimal to decisive. Results do not speak for themselves so significant attention to meet user needs is in order. This may mandate alternative reports prepared for certain key users. These could well vary in the nature of recommendations, format (executive summary, succinct report, and further details), and form (interactive workshop, electronic document, etc.).

3.8 Utilisation

Earlier sections have touched repeatedly on framing the TM to assure its utilisation. This is a multifaceted issue that typically receives far too little attention. One analyst who has become a manager advised that the normal allocation of effort ought to be adjusted from about 95% analysis and 5% delivery (presentation, communication) more toward 50–50.

The checklist of Table 27.5 synthesises findings of two of our studies concerning the utilisation of empirical technology information and analyses (see "Utilisation of Empirical Technology Analyses" at http://tpac.gatech.edu).

Table 27.5. TM Utilisation Checklist

- ❑ Understand user needs; share expectations for the study;
- ❑ Engage key users in TM study formulation and in analyses;
- ❑ Build organisational support for the study; understand potential opposition;
- ❑ Assure credibility of patent analysts, methods, and findings;
- ❑ Generate vivid reports with clear punch lines;
- ❑ Provide results in time for decision making;
- ❑ Check that the right content is provided;
- ❑ Communicate effectively by providing the right level of detail tailored to different users; communicate personally & interactively.

4. LOOKING FORWARD

This treatment of patent analysis for technology intelligence — 'Tech Mining' — offers a snapshot of how to generate practical knowledge from these patent abstract sets. It is necessarily limited in using a single case analysis (fuel cells), from one patent database (DWPI), with particular software (*VantagePoint*). We have mentioned in passing other patent resources, such as citations. Patent analysis must be tailored to meet user needs, which vary by issue, user background, and personal preferences. The potential is vast and applications are developing quite rapidly.

Four key features bear reflection. First, one can gain vital perspective by analyzing the totality of patents relating to a topic. 'Patent Profiling' is very different from finding particular, critical patents relating to patentability and other issues. Second, text mining software enables one to slice into those data in many ways to investigate points of concern. We illustrated a few analyses (maps, trends, etc.). Third, one can compose empirical measures for a purpose — 'innovation indicators' that help to assess the prospects of successful innovation. Indicators can be tailored to address specific technology management or policy questions. Fourth, by developing standard sets of such indicators, managers become more familiar with them. Their generation can be expedited through scripting of repetitive software steps, thus greatly speeding the technology analysis process. That enhances utility by making analyses that might once have taken months available in hours.

We have inventoried five sorts of professionals who have significant reasons to engage TM:

- Patent Providers: Patent offices and database providers are increasingly involved with how their information can be used. They can promote TM use of patent data by providing easy search and bulk retrieval of patent abstracts and citation data. Data providers are becoming involved with TM software too — e.g., MicroPatent offers *Aureka*; Derwent and Delphion offer *Derwent Analytics*. This can make for seamless linkage of data and tools to exploit them.
- Information Specialists: Technical librarians and search specialists need to adapt to TM opportunities. In particular, we see strong prospects for them to become gatekeepers, training others in how TM software can add value. We foresee information specialists increasingly also becoming TM analysts and participating as such on research teams.
- Researchers and Inventors: Technical professionals can take advantage of TM capabilities to locate their R&D in the technological landscape. This can aid in proposals, R&D project focus, and R&D prioritisation.
- Technology Analysts: Patent analysts and competitive intelligence analysts have been the leading users of TM and should continue as such.
- Technology Managers and Policy Makers: They have much to gain from having TM performed to enrich the empirical bases for decision–making.

Various organisations have different TM needs and opportunities. Our experience finds large firms most actively engaged in TM for competitive intelligence purposes. Smaller firms confront issues in accessing data and composing the requisite TM skills. They may be well served by third party provision of TM services. Public sector roles build around various technological innovation aspects. These include support of R&D operations (proposal assessment, research portfolio management, R&D evaluation), technology transfer (identification of opportunities, technology insertion, commercialisation), and so forth. In essence, TM provides intelligence on the R&D landscape to whomever can gain from this in any technology management situations.

We close with brief depictions of developments apt to extend the realm of TM:

- Empirical knowledge leads to better decisions than intuition alone; over the coming decade managers who take advantage of TM will outperform those who do not; technology management will come to require use of information resources.
- Companies are beginning to incorporate systematised analyses of R&D information into their strategic business processes; this will dramatically accelerate use of TM.
- Innovation indicators derived from TM can be tailored to address specific technology management issues effectively (e.g., IP asset management, R&D portfolio prioritisation, strategic planning, new product marketing).
- Generation of such indicators can be expedited by scripting the routine data treatment, analysis, and representation steps; this enables provision of intricate (but standardised) findings quickly (e.g., to do in a day what previously may have taken months), thereby greatly increasing demand.

So, we see strong prospects for these forms of patent analyses to inform technology management and policy analyses.

REFERENCES

Anderson, P., Tushman, M.L. (1990). Technological discontinuities and dominant designs: A cyclical model of technological change. *Administrative Science Quarterly*, 35, 604–633.

Börner, K., Chen, C., Boyack, K.W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology,* 37, 179–255.

Boyack, K. (2003). *An indicator–based characterization of the Proceedings of the National Academy of Sciences*. Paper presented at the NAS Sackler Colloquium on Mapping Knowledge Domains; also under submission to PNAS.

Campbell, D., Stanley, J. (1963*). Experimental and quasi-experimental designs for research*. Chicago, IL: Rand-McNally.

Chen, C. (2003). *Mapping scientific frontiers: The quest for knowledge visualization*. London: Springer.

Cohen, H.S., Keller, S., Streeter, D. (1979). The transfer of technology from research to development. *Research Management,* 22 (3), 11–17.

Dror, I. (1989). Technology innovation indicators. *R&D Management,* 19, 243–249.

Dunphy, S.M., Herbig, P.R., Howes, M.E.(1996). The innovation funnel. *Technological Forecasting & Social Change,* 53, 279–292.

Ernst, H., (2003). Patent information for strategic technology nanagement. *World Patent Information*, 25 (3), to appear.

Granstrand, O. (1999). *The economics and management of intellectual property*. Cheltenham, UK: Edward Elgar.

Lane, P., Makri, M. (2000). *Responding to diminishing technological opportunities: A socio-cognitive model of science and innovation*. Arizona State University, College of Business, working paper.

Linstone, H.A. (1984). *Multiple perspectives for decision making: Bridging the gap between analysis and action*. Englewood Cliffs, NJ: Prentice Hall.

Mahajan, V., Muller, E. (1996). Timing, diffusion, and substitution of successive generations of technological innovations: The IBM mainframe case. *Technological Forecasting & Social Change,* 51, 109–132.

Metcalfe, J.S. (1988). *The diffusion of innovations: An interpretive survey*. In G. Dorsi, C. Freeman, R. Nelson, G. Silverberg, L. Soete (Eds.), Technical change and economic theory. (pp. 560–589). London: Printer.

Metcalfe, J.S. (1981). Impulse and diffusion in the study of technological change. *Futures,* 13, 347–357.

Millson, M.R., Raj, S.P., Wilemon, D. (1992). A survey of major approaches for accelerating new product development. *Journal of Product Innovation Management,* 9, 53–69.

Modis, T. (1993). Technology substitutions in the computer industry. *Technological Forecasting & Social Change,* 43, 157–167.

Mogee, M.E. (2003). *Integrating patent analysis and technology intelligence techniques: A case study.* PIUG 2003 (Patent Information Users Group), Chicago, May 3–7.

Narin, F., Hamilton, K.S., Olivastro, D. (1997). The increasing linkage between U.S. technology and public science. *Research Policy,* 26, 317–220.

Porter, A.L., Kongthon, A., Lu, J-C. (2002). Research profiling: Improving the literature review. *Scientometrics*, 53, 351–370.

Porter, A.L., Cunningham, S.W. (under review). *Tech Mining.* New York: John Wiley.

Porter, M.E. (1985). *Competitive advantage: Creating and sustaining superior performance*. New York: Free Press.

Rogers, E.M. (1983). *Diffusion of innovations*. 3rd ed., New York: Free Press.

Schumpeter, J.A. (1975). *Creative destruction: From capitalism, socialism and democracy*. New York: Harper.

Small, H., Griffith, B. (1974). The structure of scientific literatures. *Science Studies,* 4, 17–40.

Smith, C.G. (1992). Understanding technology substitution: Generic types, substitution dynamics and influence strategies. *Journal of Engineering/Technology Management,* 9, 279–302.

Souder, W.E. (1987). *Managing new product innovations*. New York: Lexington Books.

Souder, W.E., Nashar, A.S., Padmanabhan, V. (1990). A guide to the best technology transfer practices. *Technology Transfer*, 5–16 (Winter-Spring).

Teichert, T., Mittermayer, M-A. (2002). Text mining for technology monitoring. *IEEE IEMC 2002*, 596–601.

Trajtenberg, M. (1990). A penny for your quotes: Patent citations and the value of innovations. *The Rand Journal of Economics,* 21 (1), 172–187.

Trippe, A. (2003). Patinformatics: Tasks to tools. *World Patent Information*, to appear.

Van Raan, A.F.J. (Ed.). (1988). *Handbook of quantitative studies of science & technology*. Dordrecht, The Netherlands: North Holland.

Watts, R.J., Porter, A.L., (1997). Innovation forecasting. *Technological Forecasting and Social Change*, 56, 25–47.

Zhu, D., Porter, A.L. (2002). Automated extraction and visualization of information for technological intelligence and forecasting. *Technological Forecasting and Social Change*, 69, 495–506.

Chapter 28

KNOWLEDGE NETWORKS FROM PATENT DATA

Methodological Issues and Research Targets

Stefano Breschi[1] and Francesco Lissoni[2]

[1] *Cespri, Dept. of Economics, Università "L. Bocconi", Milan, Italy*
E-mail: stefano.breschi@uni-bocconi.it.
[2] *Dept. of Mechanical Engineering, Università di Brescia, Brescia, Italy*

Abstract: The economic literature on technical change has increasingly relied upon patent citation data to measure inter-personal knowledge flows. Many doubts exist about whether patent citations really reflect the designated inventors' knowledge of both their technical fields, and of the other inventors and experts therein: citations, in fact, come mainly from the patent examiners, and possibly the patent applicant's lawyers, rather than from inventors themselves. Unfortunately, most of the papers dedicated to discussing these interpretation issues deal with USPTO data, whose citation rules are quite exceptional if compared to those of other patent offices. In addition some confusion exists between the two issues of *awareness* (whether citing inventors actually knew of the cited patents) and *existence* of a knowledge flow (whether some information on the contents of the cited patents has however reached the, possibly unaware, citing inventor). Questionnaires addressed to inventors are severely affected by this confusion, and can hardly dispel the existing doubts. We then propose to apply social network analysis to derive maps of social relationships between inventors, and measures of social proximity between cited and citing patents. Logit regressions demonstrate that the probability of observing a citation is positively influenced by such proximity. In order to perform such regressions, however, a specific sampling scheme has to used, which we also illustrate and discuss.

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 613-643.

1. INTRODUCTION

To say that 'networks' matter for technological innovation is nowadays almost to state the obvious. It is in fact widely recognised that both the creation and the diffusion of new ideas are processes which imply the integration and recombination of existing knowledge coming from different sources, locations and organizational positions.

Social scientists from various disciplines have evoked the concept of social networks also when dealing with the diffusion of innovation. In particular, social networks of individuals cutting across companies' boundaries and university campuses have been often held responsible for the circulation of valuable information and for filling the air (of both Marshallian districts and hi-tech clusters) with bright new ideas (for a survey: Breschi and Lissoni, 2001a,b).

Yet, as fashionable as it has now become to speak of knowledge and social networks, quantitative research on this topic is still in its infancy. In contrast with the abundance of case studies and narrative evidence, there are only a few papers, all of them quite recent, which attempt openly to detect social networks on a large scale, and explore their role in spreading knowledge across and within regions.

Formidable obstacles stand in the way of these efforts, both conceptual and empirical. On the one hand, the identification of useful data sources requires first a clarification of which circumstances can give birth to social networks dedicated to (or allowing for) knowledge diffusion. Which 'knowledge workers' are involved: scientists, technologists, or engineers? Which rules (commercial or not) do these workers follow with respect to sharing vs. withholding their assets? Which means do they use to communicate?

On the other hand, undertaking a wide ranging empirical analysis requires the collection of a large amount of 'relational' data, quite a daunting task for individual researchers, and one which requires considerable time and funds.

Progress on the conceptual side has nowadays made clear that the need for face-to-face contacts between knowledge workers is the key to explaining why social networks matter: such contacts are necessary to transfer knowledge assets that escape full codification, such as skills (which require practical demonstrations) and the highly specific technical or scientific jargon needed to engage in fruitful conversations. Much less agreement exists on the nature of those contacts: do they need to be as frequent as to require co-location of those wishing to engage in them? Or does an initial prolonged spell of frequent interaction ensure that future spot contacts (as in conferences), codified exchanges (as with e-mails and file sharing) and indirect interaction (e.g., reported conversations) will suffice to diffuse knowledge among network members?

On the empirical side some consensus has emerged on the unique role that patents, an old workhorse of innovation studies, can play once again in their newly discovered capacity of relational data, namely citations and co-authorship.

Citations run from one patent document to what patent examiners call 'prior art'. The latter include both earlier patents (from either the national office which grants the patent, or from other offices) and some scientific or grey literature.

Co-authorship refers to the possibility that a patent is either applied for by more than one individual or company (or any other organization, e.g., universities) or lists more than one individual as designated inventor[1]. More appropriately, we will talk of co-patenting by firms or individuals in the former case, and co-invention (and co-inventors) in the latter.

The chapter plan is as follows. In section 2 we discuss the use of patent citations as a measure of knowledge flows, placing most emphasis on the key technical issues one has to face when analysing inventor–to–inventor citations. In section 3 we discuss the use of co-invention data as a means of detecting social networks which may be held responsible for diffusing knowledge, as captured by citations. In section 4 we propose a test of the explanatory power of social networks of inventors. In section 5 we put forward a few conclusions and many questions for future research.

2. PATENT CITATIONS

Econometric studies of technological change have long exploited patents as indicators of innovation activity. As well explained by Griliches' (1990) classic survey, patent data are easily available, cover many countries, and are rich in technical information, thanks to their fine classification. The US Patent & Trademark Office (USPTO) and, from the 1980s, the *European Patent Office* (EPO) are the most heavily used sources[2].

[1] Legal persons can also appear as inventors, but it is an uncommon occurrence.

[2] USPTO data are now more easily accessible than ever, thanks to the publication of the NBER/Hall-Jaffe-Trajtenberg Patent Data File for 1975-1999, by Jaffe and Trajtenberg (2002). The dataset contains over 2 million USPTO patents, nearly 16 million patent citations, and more than 4 million 'raw' inventor records (Hall et al. 2001). EPO data are also on their way to become as much as useful, thanks to a number of projects sponsored by the European Commission (e.g., Breschi et al., 2003). A large number of providers also

In the past 15 years or so, traditional patent counts (and the related statistics on countries' and firms' patent shares) have been increasingly complemented with the analysis of patent citations, mainly for three purposes.

First, citations have been used along with patent re-classification and co-word analysis when searching for technology families, and for comparisons of the knowledge base of different companies (Pilkington, Dyerson and Tissier, 2002).

Second, citations have been used to prove that the quality of individual patents increases with the number of citations received. This allows evaluating the economic value of companies' patent portfolios[3].

Third, and most interesting for us, citations have been increasingly interpreted as 'paper trails' left by knowledge flowing from the inventor or applicant of the cited document to the inventor/applicant of the citing one.

Studies of the geography of knowledge spillovers stand out as pioneers of this application: by comparing the geographical location of the inventors (or the applicants) of the cited and the citing patents, one can hope to test the long standing Marshallian hypothesis about the existence of some spatial boundaries to knowledge diffusion. The classic methodology is the one proposed by Jaffe et al. (1993; from now on JTH93), who measure the probability of any two patents within the same technological field to be co-located (at the city, regional, or national level), and find that citation–linked patents are more likely to be co-located than unlinked ones. As an explanation JTH93 suggest that geographical proximity favours interpersonal knowledge flows, as captured by the citation patterns.

The main reason for looking at patent citations as useful 'flow' indicators resides in the belief that invention is a cumulative and social process. Inventors need to exchange with other scientists and technologists (amongst which one expects to find other inventors) many bits of knowledge which are not retrievable from bibliographic sources and personal experiments, since they escape full codification and need to be passed on by practical demonstrations, clarification of terminology through examples and metaphors, debugging of codified messages and so forth[4].

---

exist, which produce more technologically oriented, and very expensive, datasets and regular updates.

[3] For two representative studies: Hall et al. (2002); Sampat and Ziedonis (2002).

[4] An extensive literature, from both the sociology and the history of science and technology, insists upon the great importance of so-called tacit knowledge for the inventive process, as well as on the need of face-to-face contacts to transmit it. For a comprehensive survey see Cowan et al. (2000).

The main criticism levied against the use of patent citations as "flow" indicators derives from the remark that patent examiners, rather than inventors, are ultimately responsible for the citations attached to patent documents: independent inventions may be linked by a citation, one which is necessary for legal reasons but bears no diffusion meaning.

Even when citations track down effectively some sort of knowledge flow, it remains to discuss whether the latter runs between the inventors of the cited and the citing patent (inter-personal knowledge flow), or more simply between the cited patent and the inventor who cites it, such as when the inventor retrieves patent information directly from a database (direct retrieval).

More difficulties originate from the frequent confusion, at the conceptual level, between the two issues of the inventors' exposure to inter-personal knowledge flows and their awareness of such exposure.

At a more empirical level additional problems originate from the different patent examination procedures of the USPTO and the EPO (and other patent offices), which result in different mechanisms by which citations are added to patent documents.

In what follows we try to discuss the circumstances that make the use of patent citations acceptable, and to dispel some of the confusion.

2.1 Citation Rules

Citations are references either to previous patents (issued by the same patent office or by other offices) or other literature (mainly, scientific literature) to be found on the 'search report' attached by patent examiners to patent applications[5]. They help both the examiner and the applicant to judge the degree of novelty and the inventive step of each application. After receiving the search report, the applicant should have enough information to decide whether to go on pursuing the patent (which requires paying additional fees) or to give up, because the risk of rejection has been proved too high. Citations on the search report also form the basis for future search activities, especially by opponents wishing to challenge the patent's validity in court.

[5] Search reports by EPO examiners are separate documents one can find attached to patent applications, once published. As for USPTO, no separate 'search report' is published; however, the examiner's citations, as opposed to the applicant's, are listed separately on the front page of the patent document (Karki, 1997). More citations can be found in other sections of both the EPO and USPTO patent documents, such as those dedicated to describing the invention or the novelty claims. However, these are much less easily available in electronic format, and much more erratic in their frequency.

The USPTO requires applicants to disclose all the prior art they are aware of and deem relevant to this end ('duty of candour' rule), so we presume that many citations, although filtered by the examiner, were first proposed by the designated inventors[6]. The EPO does not impose any requirement of that kind, so that all the citations come straight away from the patent examiners[7].

The EPO places great emphasis on the thoroughness and parsimony of its 'patentability search' procedure: the examiners report only the prior art that really threatens the patentability of the invention. In contrast the USPTO provides a broader 'documentary search', aimed at collecting any reference which the applicant or the examiner suggest to be somehow useful in understanding the application contents (Akers, 2000). The following statements confirm this difference:

> "According to the EPO philosophy a good search report contains all the technically relevant information within a minimum number of citations" (Michel and Bettels, 2001, p.189).

> "[The USPTO examiner's] purpose is to identify any prior disclosures of technology … which might be similar to the claimed invention and limit the scope of patent protection … or which, generally, reveal the state of the technology to which the invention is directed" (OTAF (1976), as cited in Hall, Jaffe, and Trajtenberg, 2001, pp. 14–15).

When it comes to counting the number of citations per patent, the USPTO stands out as an exception: the average number of citations reported on its patents is much higher than similar figures for the EPO, and also for large national offices such the Japanese or the UK one[8]. In addition, Hall et

[6] Formally, USPTO applications may come only from individual inventors who assign their rights to legal persons such companies and other organizations only after the patent has been granted. So, ideally, all the prior art cited in observance of the "duty of candour" rule come from the inventors themselves. Of course this is not the case: it is the inventors' employers who actually manage the application procedure, with their legal and patent intelligence aids actually choosing the prior art to be cited (even truly independent inventors rely upon such aids). We come back to this in the remainder of the chapter.

[7] In contrast with the US legislation, the European Patent Convention allows legal persons to apply for patents, so when dealing with citations one needs to distinguish between inventors and applicants much more clearly than it happens with USPTO patents. Applicants may supply an additional list of references, which they may deem useful to assess the state of the art, possibly to influence the examiner or the counterpart in any foreseeable legal battle.

[8] According to Michel and Bettels, UPSTO patents cite on average about 13 other patents, and about 3 non-patent documents, while the same figures for EPO patents are 4 and one. For

al. (2000) make clear that some kind of 'citation inflation' phenomenon may have affected USPTO patents in recent times, owing to the booming patenting activity of US companies, which has placed an increasing burden on patent examiners[9].

In conclusion, the messages one can obtain from EPO citations are much less 'noisy' than those from the USPTO ones. With EPO patents we can safely presume that all the citations have been chosen by the examiner, no matter whether the inventors knew about them in advance. With USPTO patents, confusion reigns about who is entirely responsible for the front page citations: it is only since January 2001 that indications have become available on whether individual citations come from the examiner or the inventor[10].

In addition, cited–citing patent couples retrieved from EPO databases may be legitimately assumed to be 'closer', both in time and as for technological content, than those coming from USPTO data.

Unfortunately, most of the available methodological reflections on the value of citation data as indicators of knowledge flows come indeed from USPTO data users, who have been busy discussing problems absent from EPO data. To make things worse, that discussion has been hampered by a few hidden conceptual problems, to which we now turn to examine.

USPTO patents applied for in 1990 , Agrawal et al. (2003) calculate 10.2 average citations; our own calculations for EPO patents reveal about 2.8 citations received over 10 years of life (Breschi et al., 2003). However, when one compares the search reports issued by the USPTO and the EPO for international patent applications subject to Patent Cooperation Treaty (PCT), all of these differences disappear, with the USPTO figures converging towards EPO values. It is the “duty of candour” rule and the “documentary search strategy” which make the difference: when examining PCT patent applications, in fact, both the USTPO and the EPO have to stick to the same set of rules issued by the World Intellectual Property Organization (WIPO), and differences in the citation figures disappear.

[9] Clashing against time constraints and the USPTO rules for the 'documentary search' strategy, this burden may have forced the examiners to be less and less selective in picking up the right references to place on their reports. On the increasing patenting activity of US companies see also Kortum and Lerner (1998).

[10] Thompson (2003) finds that for a sample of about 2,600 citing patents and 31,000 citations, examiners account for more than 41 percent of the total citations. Besides, for 38 percent of the citing patents examiners are responsible for all the citations (the equivalent figure for inventors is only 8.5 percent).

2.2 Inter-personal Knowledge Flows and the Awareness Issue

When discussing the validity of patent citations as indicators of inter-personal knowledge flows between inventors, two preliminary questions stand out:

1. Did the citing inventor build upon the knowledge of the cited patent (cumulative effort), or did he produce from scratch both the contents of the cited patent and those of his own patent (duplicative effort)?
2. When the citing inventor produced his own invention, was he aware of the existence and contents of cited patent?

The two questions should be answered separately. The citing inventor may know very well some prior art details and exploit them successfully, without taking care of checking the exact sources of those notions, not even when he decides to draft the technical section of a patent application, which he then trusts to the hands of his (or his applicant's) attorney for the completion of the legal part[11]. Following the application the examiner finally digs out of the prior art the exact reference: the inventive effort is cumulative, although the inventor's awareness of the cited patent is absent.

On the other hand, the cited and the citing patents may well be two 'independent inventions': the citing inventor has involuntarily duplicated the cited inventor's research efforts, but he has discovered it too late (i.e., only after his attorney's or the examiner's patent search). In this case there is neither awareness nor intellectual debt between cited and citing patent.

Notice that we have not yet used the expression 'inter-personal knowledge flow', but only the expression 'intellectual debt'. While the latter points simply to a cumulative link between the contents of the two patents (cited and citing), the former presumes an exchange of information between inventors.

To stress the difference we may ask, when facing cumulative inventive efforts, whether the intellectual debt results from the citing inventor's direct retrieval of information from the prior art (which implies awareness), or from his exposure to some word of mouth diffusion process (which does not), or from both.

[11] The citing inventor may have relied upon some non-patent literature to obtain the knowledge assets he needed, without checking whether there was any prior patent literature. Or he may have simply relied upon a colleague's advice, draft, or notes; or even just on his own memory, which preserved the relevant technical information, but faulted him on the sources of such information.

Whenever direct retrieval is the only source of the intellectual debt, no interpersonal knowledge flow runs between the cited and the citing inventor, but the latter is well aware of the existence of the cited patent. At the opposite end, as suggested by our answer to question 1, the technical information contained in the cited patent may reach the citing inventor, whilst news about the existence of that patent may not.

Finally, direct retrieval and inter-personal knowledge flows may co-exist, as when the patent retrieval is indeed the first step of the invention process, but one which is spurred by some hints received by the cited inventor, or his social circle[12].

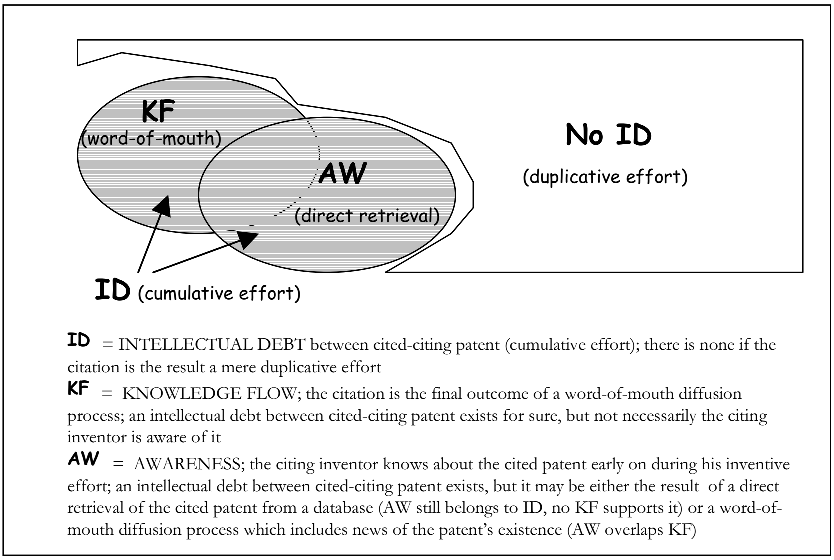


Figure 28.1. The meaning of patent citations

Figure 28.1 offers a graphic synthesis of our discussion. It stresses that many citations may signal duplicative efforts, or cumulative efforts taking place with no knowledge exchanges among inventors. At the same time, it makes clear that knowledge flows may be captured by patent citations even when inventors are unaware of those citations.

[12] Co-existence occurs also when direct retrieval kicks off the invention process, but no further progress is made until the citing inventor finds the way to retrieve some 'uncodified' bits of complementary knowledge, which again requires some kind of inter-personal knowledge flow from the cited inventor.

The size of the various sets of figure 28.1 carries no meaning, as they were chosen just to fit the page: that is, the figure is uninformative about the relative frequency distribution of the various cases. Indeed, it could hardly be otherwise, since attempts to measure those frequencies have just started, and are hampered by some confusion between the two issues of 'awareness' and 'knowledge flow'.

Two quotes from as many pioneering papers in the field may clarify what we mean by 'confusion':

> "The patent citations … have two possible sources: (a) the inventor and the patent lawyer and (b) the patent examiner. […] [Some] citations represent direct technological influences on a particular innovation, while other *citations may only represent indirect technological influences (since the patent examiner added them)*" (Almeida and Kogut, 1999, p. 908. Italics are ours).

> "[It] is likely that most citations that are not spillovers are of a different sort: *citations (added by the examiner) to previous patents of which the citing inventor was unaware. Clearly, no spillover occurs in this case"* (JTH93, p. 584. Italics are ours).

It is not clear what Almeida and Kogut mean by 'indirect' vs. 'direct' influences, but from the main theme of their paper we guess they are concerned with the possibility that two patents linked by an examiner's citation are unlikely to signal a knowledge flow between the respective inventors. Even more clearly, JTH93 rule out any knowledge flow when the citation is provided by the examiner (for the purpose of the present discussion, JTH's 93 use of the term 'spillover' is the same as our use of 'knowledge flow'[13]).

Indeed, Thompson (2003), whose data distinguish between inventors' and examiners' citations, finds that the former are more likely to show JTH-like co-localization effects. This encourages us to presume the existence of

[13] We prefer writing about 'knowledge flows' because this term bears no implication with respect to the intellectual property regime that rules the exchange of information between inventors. Spillovers are externalities, which means that the citing inventor has not paid for the information he has received. In fact, when discussing citations between patents from the same company, JTH93 talk of 'internalised spillover', to stress the existence of well established IPRs over the exchanged information. However, as hinted by Zucker et al. (1998), and proved by Møen (2000), it may also be that citations running across firms are the result of implicit contractual arrangements, by which either the cited inventor or his employer (i.e., the patent grantee) get paid for the information delivered.

some interpersonal knowledge flow. However, Thompson finds that also examiners' citations exhibit non-negligible co-location effects.

From our viewpoint, therefore, there is no reason to exclude that examiner's citations (i.e. 'unaware citations') may signal a knowledge flow. At most we can presume that the cited and the citing inventors do not know each other, otherwise the former would have passed both the information on the patent's contents and the information on the existence of the patent itself. But it may well be the case that the two are linked by one common acquaintance, or a chain of social acquaintances, who are responsible for passing on the information on the patent contents, as well as the necessary 'tacit bits' to build upon them: the longer the chain, the more likely the case that other bits of information, specifically those on the existence of the cited patent, have been lost during the diffusion process. Here again, Thompson's (2003, pp. 6–7) study is revealing, when it shows that inventors are hardly better than examiners at retrieving their own company's prior art, which 'suggests at the least a certain casualness with which inventors prepare their applications'.

Notice that the literature we quoted is somehow uncertain when it comes to accepting that knowledge may flow from the cited to the citing inventor not just through direct communication (the two inventors know each other), but also via a social chain of personal relationships. Notions from social network analysis, and in particular from recent advancements discussing the length of social chains within large communities of scientists, are absent from many of the most influential papers exploiting the relational content of patent citation data[14].

2.3 Testing the Knowledge Flow Hypothesis: Questionnaires versus Co-authorship

Finding any evidence of the connection between knowledge flows and patent citations is very hard. Early suggestions from the sociology of science point out that independent inventions are very likely to arise both within and across different communities of scientists and technologists. However, none of the evidence from the classical literature on independent inventions is based on patents; rather, it exploits historical records such as biographies and

[14] Modern social network analysis applied to innovation diffusion studies dates back at least to Valente (1990). Emphasis on social closeness within a large population has been placed by so-called "small world" theories, of which the most up-to-date and at the same time most readable survey has been produced by Watts (2003). None amongst the papers we have quoted so far ever mentions such theories, nor the more basic concept of 'social network'.

ancient chronicles, and compares, in many cases, different epochs (Ogburn and Thomas, 1922; Merton, 1961).

To our knowledge the only patent–based study which has investigated the meaning of citations at the inventor level has been conducted by Jaffe et al. (2000a). They interviewed more than 150 inventors in order to

> "[…] try to learn about the extent and modes of their communication with earlier inventors, and about the extent to which the appearance of citations in their patents is indicative of this communication" [15].

When asked about what spurred their research efforts, the interviewees assign the same weight to word–of–mouth and direct queries of the patent/technical literature. But they attach a much greater importance to some kind of 'awareness of a commercial opportunity', an answer which goes hand in hand with the frequent 'unawareness of the cited patent'.

Once again, Jaffe et al. (2000a) interpret unawareness as 'absence of spillovers (knowledge flows)'. This is questionable. To see why, notice that when asked about the 'technology underlying the cited patent', more than 40% of citing inventors suggest they knew about it before or while working on their own invention, about 25% answer that they discovered it after developing the invention, and only 25% observe that they were totally unaware of it until reached by the questionnaire. In contrast many less than 40% of the inventors knew about the cited patent itself, and many more than 25% became aware of it during the development process (which means they may have been influenced by it). That is, technology awareness and patent awareness do not go hand in hand.

Even more significantly, only 6% of the cited inventors reported some direct communication with the citing one, whilst no less than 75% admitted to know of the citing inventor or at least of his research. Such contrast suggests the existence of a social chain linking citing and cited inventors, one which provides a useful word–of–mouth information channel even in the absence of direct communication.

More importantly, it is questionable whether interviews and questionnaires addressed to inventors alone may help in dispelling all the existing doubts: awareness problems may plague not only the inventors' recollection of their knowledge of prior art, but also their recognition of the intellectual debt they owe to their colleagues or the prior art[16].

[15] Jaffe et al. (2000a: 1). Citations considered by Jaffe and his colleagues exclude so-called "self-citations", i.e. citations linking two patents owned by the same grantee (see below). For further discussion of the survey's results see also Jaffe et al. (2000b).

[16] Jaffe and his co-authors find that cited and citing inventors disagree about the knowledge debt of the citing patent towards the cited one, which again proves how difficult it is to

An alternative strategy for testing the validity of patent citations as knowledge flow indicators consists in recovering citation–independent relational data on the inventors' social networks, and then use them to test how 'social proximity' can help interpreting the observed citation patterns. In section 3 we discuss that strategy.

2.4 Self-Citations: Definitions and Technical Problems

All the existing studies making use of patent citations go to a great length in discussing the methodological problems raised by so-called 'self-citations'. These are generally intended as citations running across patents from the same applicant. Whenever citations are used to assess the value of the cited patent, self-citations have to be excluded to correct for the bias of the applicant's patenting strategy (which may be biased towards 'thin' patents, each of them carrying a few claims, but all of them related to each other). As for studies of citations as knowledge flows, concern about the need to distinguish 'pure' spillovers from internalised ones makes it necessary to tell inter-firm citations apart from intra-firm citations.

The distinction is not an easy one. The so-called 'backward lag' (the difference between the publication years of the citing and the cited patent) is often a considerable one. In order to detect all self-citations we should be able to track all the change of properties undergone by the applicants: although the cited and the citing companies may be independent when the cited patent was issued, this may no longer be true when the citation occurs, because of mergers and acquisitions[17].

Notice that, once again, USPTO data suffer more of the drawback, since on average USPTO citations refer to older patents than EPO ones. In fact, Hall, Jaffe and Trajtenberg (2000) find for the USPTO data that 50% of all citations are made to patents that are 10 years older than the citing ones, as opposed to the 3–year figure calculated by Breschi et al. (2003).

When we move on to consider the inventor–based dataset, a further category of self-citations emerge, which the available literature usually define as 'personal self-citations'. These occur whenever the cited and the citing patents share at least one inventor. In this case we can be sure of both the existence of a knowledge flow and full awareness.

---

disentangle individual contributions to ideas. Similarly, the inventors from the University of Pavia (Italy), when interviewed by Laboranti (2004) found it difficult even to tell apart their own contribution to the patented invention from the contribution of their co-inventors.

[17] Conversely, the citing company may be a spin-off from the cited company, that is, the two companies were the same when the cited patent was issued, but have become independent by the time the citation occurs.

Personal self-citations which do not correspond to self-citations at the company level are especially interesting: they signal the potential of cross-firm inventors' mobility as a means of knowledge transfer, one which has attracted increasing attention and proved to be very relevant (see sections 3.1 and 4 below).

3. CO-INVENTION AND SOCIAL NETWORKS

Co-authorship data from scientific papers have always been a powerful tool for empirical analysis in the sociology of science, and have more recently proved useful for testing 'small world' theories on the positive relationship between the size of social networks, and the closeness of individuals therein (see again footnote 15).

The main reason for this success is that it is widely acknowledged that scientists form quite a close community, and whose distinctive codes of practice set them apart from the rest of society. Scientific ideas originate, circulate, and are improved mainly within a set of connected cliques sharing some inaccessible jargon and a very odd reward system: when studying their history or sociology we can safely assume that the most relevant social links are those within the community, concentrate on them, and set aside links towards society at large [18].

Co-authors of scientific papers rely upon mutual understanding or at least upon knowledge complementarity. We may presume that co-authors know each other so well that they can effectively exchange important knowledge assets, especially if directly relevant to the contents of their publications. Each time two scientists work on a joint paper, we can safely treat them as two nodes of a social network connected by a bi-directional link (which can also be weighted by considering whether other joint papers exist, and their scientific relevance). By considering all scientists within a given discipline, we can build the entire social network for that discipline, and proceed to explore its structural properties (very much along the lines of Newman, 2000 and 2001)[19].

[18] Inaccessibility of much of the scientific jargon is self-evident. As for the reward system, we refer here to Merton's definition of science as being driven by an "institutionalised system of open-communication-and-correlative-reward" (Merton, 1977, p. 48).

[19] For a long time citations to and from scientific papers have been exploited as a useful source of relational data. Scientists cite each other for a number of reasons, the main one being acknowledging other scientists' priority or authority. This allows us to compare two scientists' citation set, and uncover some common roots. Patent citations serve much less effectively the same purpose, for all the reasons we outlined in section 2. Notice also that

We propose here to extend the same methodology to co-authorship of patents, in particular to co-invention (as opposed to co-patenting; see section 1 for definitions). We argue that, at least within some technological fields, inventors' communities may be as close and as self–referential as it is necessary to deal with them without considering knowledge inputs from the outside (or, better, without considering the exclusion of outside inputs from the study as too great a damage to the study itself). We find support to this analogy in Constant (1984).

3.1 Citations and Inventor Datasets: a New Direction

If we assume that inventors listed on the same patent know each other, and have possibly exchanged key technical information, classifying patents by inventor becomes a crucial scientific exercise. To our knowledge, extensive efforts in that direction have just started, and only a handful of studies are available[20].

Agrawal et al. (2003) consider about 50,000 USPTO 'originating' patents by North American companies. Each patent is assigned to ('located in') a metropolitan area where the majority of inventors reside. After excluding personal self-citations Agrawal et al. move on to check whether citations to the 'originating' patents come preferably from co-located patents or from patents located elsewhere, but whose inventors used to work (i.e., signed their first applications) in the same metropolitan area as the originating patent's.

Finding evidence that not only current geographical co-location matters (as in JTH93), but also that previous co-location does, the authors conclude that "one plausible interpretation of this finding is that individuals invested in social ties with others at their prior location during their residency there, and at least part of their social capital endured to support above average knowledge flows back to their prior location (op. cit., p. 2)".

Further indirect evidence comes from studies of the role of inventors' mobility across firms in spreading knowledge. Song et al. (2003) set up a

---

patent citation show statistical patterns which are much different from those of scientific citations, first and foremost by not displaying any sign of being affected by the so-called 'Matthew effect' (see data reported by Karki, 1997). This is understandable, since no reward system (neither for the inventors, nor for the applicants) is attached to patent citations, certainly no one as sophisticated and decisive for individuals' careers as the one attached to citations of scientific papers (Merton, 1988).

[20] Fairness requires pointing to Sirilli (1987) as the true pioneer study using an extensive inventor-based patent data set. However, Sirilli's data set was never regarded, not even by its author, as a source of relational data, nor was it ever coupled to citation data.

database on 180 'mobile' engineers, who signed about 500 patents for a number of US semiconductor companies. For each engineer they consider his first patent in time, and test whether companies hiring him at a later date did or did not cite that patent. They find that the citation probability increases when the hiring firm is taking patents outside its traditional core technologies, while geography (here represented by the 'domestic vs. international' dichotomy) bears no consequences.

3.2 From Co-invention to Social Networks

Other authors have made bolder efforts in the direction of explaining citation patterns by applying social network concepts. Stolpe (2002) set up a data set of more than 2,000 inventors and about 1,300 patents in the field of Liquid Crystal Display technology, with information about the inventors' patterns of mobility across the 200 patent assignees. He then proceeds to test whether those mobility patterns can explain the citation links across firms: starting from a sample of originating patents, he builds both the related set of citing patents, and an 'endogenously stratified' control sample of non-citing patents, which match closely the technological class of the citing ones[21].

A logit regression is then run to find the determinants of the citation probability (excluding personal self-citations[22]). Amongst such determinants, Stolpe includes the existence of a 'prior cooperation' link between the inventors of key patents and those of the citing and non-citing patents, but with negative results: prior cooperation between inventors is found to hold no explanatory power. We observe, however, that such an explanatory variable is still quite a rough proxy for social distance, one which points

[21] Endogenous stratification techniques help solving a major technical problem. Citations are a relatively rare event; whenever the patent database is large enough to generate a large number of citations, the resulting number of patent pairs is no more manageable from the computational viewpoint. Endogenous stratification consists in collecting all observations for which the dependent variable takes value 1 (a citation exists; we define them as the 'cases'), but only a random selection of *n* observations for which the dependent variable takes value 0 (no citation links the pair of patents; we refer to them as the 'controls'). Albeit with slight differences, this sampling procedure has been recently adopted by other authors dealing with citation data, such as Sorenson and Fleming (2001) and Sorenson (2003). For more methodological remarks on this strategy, which applies to all instances of 'rare events' (of which citations are a chief example) see also King and Zeng (2001). For an application, see section 4 below.

[22] The exclusion of personal self-citations is motivated by Stolpe's exclusive interests in pure spillovers, rather than generic knowledge flows, under the assumption that personal self-citations signal the inventor's capability of fully appropriating the returns from his invention.

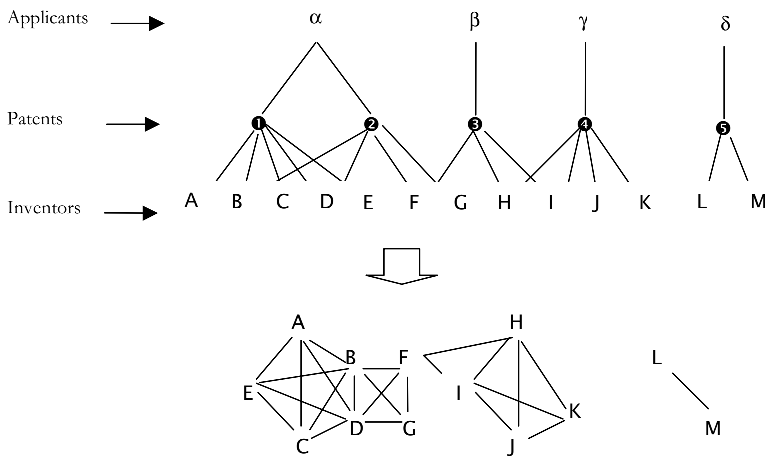
exclusively on the inventors' mutual acquaintance, and excludes longer social chains[23].

Our own research efforts have been directed at overcoming this limitation, by the creation of the EP-INV data set, which contains all EPO applications signed by Italian inventors from 1978 to 1999, and report 'clean–and–checked' data for 30,243 inventors (name, surname, address) and 38,868 patents (both granted and not granted). The dataset allows us to detect all the social connections among inventors, as originated from previous co-invention experiences, from whatever technological field. We presume that two inventors who worked together on at least one patent will keep in touch after then, or will anyway be capable of getting in touch again to exchange information or share some knowledge assets. Co-invention data can then be exploited to map the complex web of social ties among inventors, and measure a number of 'structural properties' of such web, typical of social network analysis (Wasserman and Faust, 1994).

The following hypothetical example, taken from Balconi et al. (2004), illustrates the main idea (see Figure 28.2). Let us suppose we face five patent applications (1 to 5), coming from four different applicants (α, β, γ, δ). Applicant α is responsible for two applications (1, 2), whilst applicants β, γ, and δ, one each. Patents have been produced by thirteen distinct inventors (A to M). We can reasonably assume that, owing to the collaboration in a common research project, the five inventors are 'linked' to each other by some kind of knowledge relation. The existence of such a linkage can be graphically represented by drawing an undirected arrow between each pair of inventors, as in the bottom part of figure 28.2. Repeating the same exercise for each team of inventors, we end up with a map representing the network of linkages among all inventors[24].

[23] In addition, Stolpe considers only co-operations having taken place in some third organization. That is, he does not consider all cases in which the cited and citing inventors used to work together either for the assignee of the cited patent or for the assignee of the citing/control patent. Again, exclusive concern for pure spillovers explains this odd choice.

[24] In the language of graph theory the top part of the figure reports the affiliation network of patents, applicants and inventors. An affiliation network is a network in which actors (inventors) are joined together by common membership in groups of some kind (patents). Affiliation networks can be represented as a graph consisting of two kinds of vertices, one representing the actors (e.g., inventors) and the other the groups (e.g., patents). In order to analyse the patterns of relations between actors, however, affiliation networks are often represented simply as unipartite (or one-mode) graphs of actors joined by undirected edges two inventors who participated in the same patent, in our case, being connected by an edge (see bottom part of Figure 28.1).



Top: Bipartite graph of applicants (α, β, γ, δ), patents (1 to 5) and inventors (A to M), with lines linking each patent to the respective inventors and applicants.
Bottom: the one-mode projection of the same network onto just inventors

Figure 28.2. Bipartite graph of patents and inventors (Balconi et al., 2004)

Using the graph just described, we can derive various measures of social distance among inventors.

1. *Connectedness*. Inventors may belong to the same component or they may be located in disconnected components. A component of a graph is defined as a subset of the entire graph, such that all nodes included in the subset are connected through some path[25]. In Figure 28.2, for example, inventors A to K belong to the same component, whereas inventors L and M belong to a different component.
2. *Geodesic distance*. The geodesic distance is defined as the minimum number of steps (or, more formally, 'edges') that separate two distinct inventors in the network. In Figure 28.2, for example, inventors A and C have geodesic distance equal to 1, whereas inventors A and H have distance 3. This means that the linkage between them is mediated by two other actors (i.e. B and F). In other terms, even though inventor A does not know inventor H directly, she *knows who* (inventor B) knows who

[25] More precisely, a component of a graph is a subset of nodes, for which one can find a path between all pairs of nodes within the subset, but no paths towards the nodes outside. In our specific context a node must be interpreted as an individual scientist/inventor.

(inventor F) knows inventor H directly. The geodesic distance between a pair of inventors belonging to two distinct components is equal to infinity (there is no path connecting the two inventors).

Some inventors stand out for the number of links they exhibit, that is they exhibit high 'degree centrality', whilst others may have a particularly important role in connecting different components (they can be either by 'mobile' inventors, that is industrial researchers moving across firms, or, alternatively, by free lance or academic researchers with multiple ties to industry).

By taking any pair of patents we can examine the social distance between the respective inventors, and take the lowest possible value whenever at least one of the two patents is the result of a team effort (the distance drops to zero if one inventors is responsible, alone or with others, of both patents). For all pairs of patents with different application dates we can check whether a citation link exists, and associate the probability of its existence to the social distance between the inventors. In Breschi and Lissoni (2004) we have employed social distance measures to test the robustness of JTH93's findings on the role of geographical distance in knowledge diffusion. We found that, in the absence of social connectedness, geographical proximity can hardly explain citation patterns; in contrast, social connectedness enhances the role of geographical proximity, especially when the social distance between inventors is short. However, we did not test directly for the impact of social distance on the probability of a citation link to arise.

In that respect some useful methodological hints come from Singh (2003), who has exploited a large sample of USPTO data to build an extensive social network of the kind in Figure 28.2, albeit one in which nodes are not individual inventors but teams of inventors, and the ties between two nodes result from the existence of at least one inventor working, or having worked, with both teams.

Singh investigates directly the role of social networks in explaining citation patterns by running a regression in which the dependent variable is the presence/absence of a citation link between all possible pairs of patents (with different application dates) within his dataset, and the chief independent variable is a set of five dummies representing as many levels of social distance between the two patents. In what follows we do the same by exploiting, once again, the EP-INV database. Before doing so, however, we discuss briefly one of the chief objections to the use of inventor data for the detection of social networks supporting knowledge diffusion.

3.3 Inventors as a Relevant Community of Experts

A frequent objection raised against considering co-invention data as a rich enough mine for the extraction of relational information is that interpersonal exchanges between inventors are no more than a tiny subset of all the exchanges enabling inventors to achieve their results: personal contacts with non-inventors (such as academic researchers and non-patenting technologists) would be as important in diffusing knowledge relevant for the invention.

We suggest that, at least for a number of technological fields, this objection may not be as robust as it seems at first glance. A good point of departure is the discussion of a legal technicality which goes under the name of 'disclosure rule'.

Modern patent systems draw their justification from a basic economic trade-off: the patent assignee is granted a temporary monopoly power over an invention only if he fully describes it in the application (discloses), so that when the patent expires imitation can take place and competition be restored. This basic principle is then translated into a set of rules, which do not differ much across the various national IPR laws and international conventions (see Akers, 1999, pp. 161–162).

All the disclosure rules, in fact, point invariably to the 'average expert's understanding' as the yardstick against which one should judge whether the level of disclosure is satisfactory: that is, disclosure is sufficient ('enabling') when the invention description is clear enough for the invention to be carried out by a person "skilled in the art" (as in the EPO Guidelines for Examination; Part C, Chapter II, Section 4.18).

This means that inventors (or, better, their legal aids), when drafting a patent document, bear in mind a well defined reader's profile. In turn, such profile reminds one of a community of experts, whose core includes the inventors themselves, their colleagues (within and outside the organization they work for), and the patent examiners; anybody whose competency is not up to the profile is excluded.

An understanding of the basic scientific principles behind the technological field (such as the one mastered by academic scientists) or some knowledge of the prior art (which we can attribute to 'non-inventing' technologists) are not enough to be judged as 'average experts of the field'. More is required, namely, an active involvement in the invention process or in the technical support of the creation of intellectual property rights[26].

[26] By 'technical support to the creation of IPRs we mean the activities of patent examiners and IPR consultants with a strong technical background. For a recent study which shows how

In addition, patent 'fields' are defined quite narrowly: EPO patents rely mainly on the International Patent Classification identifies up to 628 so-called 'subclasses' and approximately 69,000 technological 'groups'[27]. According to our interviews, EPO patent examiners often specialise at the subclass level, and hardly any inventor can be found who sign patents across different classes.

Even the simpler USPTO classification is detailed enough to suggest that inventors, and possibly examiners, are highly specialised and can hardly exchange useful knowledge assets across different classes (see remarks on the JTH93 experiment by Thompson and Fox-Keans, 2003).

These observations match suggestions from the economics and sociology of innovation, according to which technical knowledge is highly specific and 'local' (see again Constant, 1984). It follows that the population of inventors is more than a tiny and unchecked sample of all the individuals who can influence inventors themselves. Rather, it represents the most immediate and influential social environment from which inventors draw ideas and information, at least for the technical contents of their patents[28].

We expect this representativeness to be the stronger, the more representative patents are of inventive activity. Fields such as Chemicals and Electronics, in which a few companies' large R&D teams produce patents at a constant rate, host for a certain number of inventors who can be regarded, and regard themselves, as a 'community of experts'. The opposite may hold for inventors in the Mechanical fields or in more traditional technologies, who are more likely to patent on an occasional basis and less likely to work in teams.

---

networks of scientists and networks of inventors hardly overlap, even in a typical science-based technology such as biomedical engineering, see Murray (2002).

[27] The IPC is based on an international multi lateral treaty administered by WIPO, the World Intellectual Property Organization, and entered into force in 1975. The industrial property offices of more than 90 States (amongst which the USPTO), four regional offices (including the EPO) and the International Bureau of WIPO under the Patent Cooperation Treaty (PCT) currently use it. The IPC is a 12 digit hierarchical classification system comprising *sections* (1 digit), *classes* (2 digit), *subclasses* (4 digit) and *groups* (*main* groups: 5-7 digit; and *sub*groups: up to 12 digit). The seventh edition has been in force since 2000 (http://www.wipo.org/classifications/en/).

[28] Sure enough, the "awareness of commercial opportunity" mentioned by Jaffe et al. (2000a; see section 2.1 above) may come from managers and entrepreneurs, but it is not this kind of influence we are concerned with here.

4. A REGRESSION TEST OF THE SOCIAL NETWORK HYPOTHESIS

In this section we propose a test of the hypothesis that direct and indirect social ties amongst individuals are a key determinant of patent citations. More specifically, indicating by P(K, k) the probability that a patent K cites a patent k, we aim to test whether such a probability increases with the degree of social connection between the two patents, after controlling for other possible factors affecting the probability of a citation. We describe first the sampling design of our experiment and then we turn to describing the variables used and the results of our estimates.

Sampling design

In principle one could approach the problem of estimating the factors which affect the probability of a citation tie between patents by looking at all possible pairs of potentially citeable and potentially citing patents (K,k), using then logistic regression to estimate the effects of covariates. Even for a relatively small country in terms of patenting such as Italy, this would require dealing with a very large data matrix, so we follow Stolpe's (2002) and Singh's (2003) in adopting an endogenous stratification sampling strategy (see again footnote 21).

In particular, we have followed a 4-step procedure:

1. We select a cohort of originating patents, e.g., 1987, by application date. Let k_{tj} be k_{th} patent in cohort t with technology class j[29];
2. For each subsequent cohort of patents, e.g., 1990, by application date, we generate all potential pairs between them and the originating patents. Let K_{Th} be the K_{th} patent in cohort $T>t$ with technology code k; the pair (k_{tj},K_{Th}) identifies a potential citation from patent K to patent k[30]*;*
3. From the set of potential citations generated in this way, we select all citations (the 'cases');

[29] In this paper we use a concordance table which maps the International Patent Classification codes into 30 exhaustive technological classes. The concordance table was originally produced jointly by FhG-ISI and OST for an earlier release of the IPC, and we have updated to the current one. It is available online at: http://www.cespri.it/srv_area/ docs.htm

[30] Let n_{tj} be the number of patents in cohort t with technology code j and N_{Th} the number of patents in cohort T with technology code h. The number of possible pairs (i.e. potential citations) is therefore given by: $\sum_{j=1}^{30}\sum_{h=1}^{30} n_{tj} \cdot N_{Th}$

4. For each citation, we select two 'control' pairs[31], such that the selected patents in each pair belong to the same 'technology cell' as the original citation, i.e., had the same respective technology classes as the patents in the original citation. The reason for stratifying the sampling procedure according to the technological classes of cited and citing patents is that the technological relatedness of two patents is a key variable affecting the probability of a citation and a simple random sampling would fail to take this into account[32].

This procedure has been applied to three cohorts of patents: 1987, 1988 and 1989. For this paper we have considered only patent applications in which the applicant was an Italian organization, excluding non-Italian applicants and individual inventors. The final sample consists of 4803 observations, of which 1,601 are actual citations and 3,202 are controls[33].

Social distance and control variables

Given a pair of patents (k_{tj}, K_{Th}) the 'social distance' between them at time (T-1), i.e., the period just before the potential citation, is defined as the lowest among the geodesic paths connecting the inventors of the two teams of inventors. The 'social distance' variable defined in this way has therefore the characteristics of a categorical variable and for estimation purposes it is convenient to transform it into a set of dummies. In particular, we define nine of them, mutually exclusive and exhausting the social distance possibilities:

a) d_0: this variable takes the value 1 for *personal self-citations*, and 0 otherwise (geodesic distance is zero);
b) d_1: this variable takes the value 1 for *prior collaborations*, namely, pairs of patents where one or more inventors of the citing patent have previously co-invented with one or more inventors of the cited patent, i.e., the geodesic distance between them is 1;
c) d_2: this variable takes the value 1 for *common acquaintances*, namely pairs of patents in which one or more inventor of the citing patent and one or more inventors of the cited patent had one (or more) co-inventors in common, i.e., the geodesic distance is 2;

[31] We decided to sample two controls to ensure that the 'control' group has roughly the same size as the set of realised citations.

[32] Steps 2 to 4 described in the text are obviously repeated several times, one for each cohort of future patents up to 1998.

[33] The 1,601 citations correspond to 969 cited patents.

d) d_{3}–d_{6} these variables take the value 1 if the geodesic distance between the patents is, respectively, 3, 4, 5, and 6, and they take the value 0 otherwise;
e) $d_{>6}$: this variable take the value 1 if the geodesic distance between the patents is greater than 6, but still finite, i.e., the inventors of the two patents belong to the same connected component in the co-invention graph[34];
f) *disconnected*: this variable takes the value 1 if the inventors of the two patents are not reachable because they belong to disconnected components, i.e., the social distance between them is infinity; we also define a complementary dummy variable — *connected* — to identify pairs of patents whose inventors belong to the same connected component in the co-invention graph.

In the estimations reported below, the reference group is disconnected, i.e., the pairs of patents whose respective inventors are not reachable.

In addition to the effect of social distance we have introduced a number of variables, designed to control for some basic factors which could affect the probability of a citation tie between two patents. First of all, we control for the effects associated with time by including two sets of control variables: the time lag (expressed in months) between the application dates of the two patents; and a set of fixed effects for the application year of the citing patents. Second, we introduce a set of controls for the degree of technological relatedness between patents. In particular, we define the following dummy variables:

a) *samesub*: it takes the value 1 if the two patents' IPC codes are in the same subclass, 0 otherwise;
b) *samegroup*: it takes the value 1 if the two patents' IPC codes are in the same group, 0 otherwise;
c) *sameclsing*: it takes the value 1 if the primary IPC subclass of the citing patent is the same as one of the secondary IPC subclasses of the patent cited, 0 otherwise;
d) *sameclsed*: it takes the value 1 if the primary IPC subclass of the cited patent is the same as one of the secondary IPC subclasses of the citing patent, 0 otherwise;

In addition to this, in order to control for the different propensities towards citing across technological fields we introduce a set of fixed effects for the technological classes of the citing patents.

[34] It should be noted that the largest, yet finite, distance between pairs of patents is 11 for actual citations and 19 for controls.

We also control for the geographical localization of the inventors who have produced the two patents by looking at the inventors' address. For each pair of patents we look at all possible dyads among inventors in the two teams and define a dummy variable (SameLLS) which takes the value 1 if at least two inventors are co-located in the same geographical area, and 0 otherwise. The spatial units of observation are the so-called Local Labour Systems (LLS), a set of functional regions defined by the Italian Institute of Statistics on the basis on employment data[35].

Finally, we make use of a dummy variable also for controlling for self-citations at the company level (same assignee).

Estimation and results

Given the choice–based sampling procedure used we estimated the effect of social distance and other controls on the probability of a citation by using a weighted exogenous sampling maximum likelihood (WESML) procedure. A weight of 1 has been assigned to all observations corresponding to actual citations, because all of them have been sampled. Observations corresponding to controls have been weighted as the inverse of the fraction of all patents, with a specific combination of technological codes, included in the sample. As the coefficients of logit estimates do not have a direct economic interpretation, in the following tables we have reported odds-ratios[36].

Table 28.1 reports estimation results for a simple model in which social connectedness has been treated as a binary variable. The first two columns refer to estimates which include company self-citations, whereas the last two columns exclude them. Results reported in columns (a) and (c) suggest that both social and geographical proximity affect the probability of a citation tie between two patents. All else equal, socially connected patent pairs are 3 times more likely to result in a citation than patent pairs whose inventors are not socially linked. Similarly, spatially co-located patent pairs are 4.6 times more likely to result in a citation than geographically separated patent pairs. It is worth stressing that technological relatedness is a major factor affecting the likelihood of a citation tie. Patent pairs sharing the same primary technological code at the subgroup level are more than 100 times more likely to result in a citation than patent pairs not sharing the same primary technological code. Moreover, confirming the results of Thompson and Fox-

[35] A LLS consists of a group of cities and towns with a self-contained labour system: the labour force residing within each LLS is characterised by high internal mobility and virtually no mobility outside the system. According to Sforzi (1991) there are 784 Local Labour Systems in Italy.

[36] This is easily obtained by exponentiating logit coefficients.

Kean (2003), our findings suggest that in order to assess the degree of technological similarity between patent pairs one should not only look at primary IPC codes, but also take into account secondary classes.

Table 28.1. Social connectedness and patent citations

	Self-citations included		*Self-citations excluded*	
	(a)	(b)	(c)	(d)
Connected	3.037[a]		17.256[a]	
	(0.694)		(5.612)	
SameLLS	4.679[a]		5.154[a]	
	(1.012)		(1.062)	
Connected*SameLLS		14.605[a]		91.258[a]
		(4.037)		(31.745)
Disconn.*SameLLS		5.322[a]		7.901[a]
		(1.429)		(2.255)
Connected*(1–SameLLS)		3.480[a]		27.945[a]
		(1.241)		(9.999)
Samesub	1.464	1.449	1.512	1.462
	(0.465)	(0.461)	(0.525)	(0.512)
Samegroup	121.38[a]	125.66[a]	200.66[a]	222.34[a]
	(36.935)	(39.105)	(72.843)	(82.151)
Sameclsing	7.210[a]	7.232[a]	3.673[a]	3.831[a]
	(2.175)	(2.173)	(1.308)	(1.349)
Sameclsed	1.079	1.069	1.328	1.303
	(0.277)	(0.274)	(0.419)	(0.407)
Same assignee	36.343[a]	35.229[a]		
	(9.405)	(9.342)		
Number of obs.	4803	4803	3891	3891
Log-likelihood	-7.698	-7.697	-3.914	-3.907
Pseudo-R2	0.404	0.405	0.308	0.310

The table reports odds-ratios, rather than logit coefficients. Robust standard errors in parentheses. The estimations include a set of fixed effects for the year of citing patents, a set of fixed effects for the technological classes of citing patents, and the time lag in months between the application dates of the citing and cited patents. The coefficients of these variables are not reported for brevity. In columns (b) and (d) the reference group for comparison is the case of patent pairs not socially connected and not geographically co-located.
[a] significant at the 1% level; [b]significant at the 5% level.

Columns (b) and (d) add to the previous analysis the interaction effects between social connectedness and geographical proximity. Results seem to suggest that, even if spatial proximity is an important mediating factor, spatial co-location is not a necessary condition for the effective transfer of knowledge. Being part of a socially connected group of individuals may indeed help to overcome geographical boundaries. For example, consider a pair of spatially separated patents: if socially connected, the pair's citation odds are 30 times larger than if socially disconnected. In contrast, for socially disconnected patents, co-location multiplies the citation odds only by 9.

Table 28.2. Social distance and patent citations: Odds Ratios

	Self-citations included	*Self-citations excluded*
d_0	35.629[a] (12.730)	495.46[a] (274.65)
d_1	21.881[a] (8.267)	149.35[a] (73.187)
d_2	3.688[a] (1.712)	16.369[a] (8.379)
d_3	5.484[a] (2.528)	8.488[a] (4.954)
d_4	0.495 (0.364)	3.444[b] (1.931)
d_5	0.611 (0.306)	0.797 (0.497)
d_6	0.265[b] (0.183)	0.541 (0.390)
>d_6 (but finite)	0.125[a] (0.076)	0.471 (0.277)
SameLLS	4.321[a] (0.964)	4.444[a] (1.205)
Samesub	1.069 (0.384)	2.209[b] (0.829)
Samegroup	128.12[a] (42.482)	140.66[a] (55.489)
Sameclsing	5.209[a] (1.576)	3.530[a] (1.316)
Sameclsed	1.913[b] (0.511)	1.132 (0.370)
Same assignee	12.650[a] (3.676)	
Number of obs.	4803	3891
Log-likelihood	-7.109	-3.554
Pseudo-R2	0.450	0.372

For explanatory notes see the legend of Table 28.1. The baseline category left out is represented by patent pairs not socially connected (disconnected).
[a] significant at the 1% level; [b] significant at the 5% level.

To move beyond consideration of social connectedness simply as a binary variable, we run a regression which explicitly incorporate the effect of geodesic distance by including all the nine dummies listed above. Table 28.2 shows that the citation probability falls quite sharply with social distance, starting from very high levels at low social distances. At distances greater than 3, the citation odds are not significantly higher than for disconnected patents; and at distances larger than or equal to 5 the odds are even significantly lower. Whilst extremely important for transmitting knowledge, the effectiveness of social connections seems to decay very rapidly with social distance. Alternatively, we can presume that long–distance links decay rapidly over time, and do not convey anymore any knowledge flow.

Table 28.3. Social and geographical proximity and patent citations: Odds Ratios

	Self-citations included	*Self-citations excluded*
d_0 * SameLLS	219.15[a] (85.440)	3555.31[a] (2244.45)
d_0 * (1- SameLLS)	24.389[a] (16.893)	365.03[a] (455.80)
d_1 * SameLLS	98.639[a] (40.290)	450.02[a] (235.41)
d_1 * (1- SameLLS)	37.302[a] (25.766)	4912.18[a] (3916.45)
d_2 * SameLLS	18.730[a] (10.052)	106.69[a] (65.410)
d_2 * (1- SameLLS)	5.569[b] (4.405)	4.688 (4.987)
d_3 * SameLLS	19.713[a] (11.756)	26.548[a] (16.599)
d_3 * (1- SameLLS)	12.149[a] (7.543)	47.968[a] (50.248)
d_4 * SameLLS	1.297 (1.195)	20.136[a] (13.867)
d_4 * (1- SameLLS)	2.938[b] (1.619)	2.803 (2.219)
Disconn.*SameLLS	6.255[a] (1.602)	7.754[a] (2.112)
Samesub	1.000 (0.362)	2.593[a] (0.952)
Samegroup	140.88[a] (48.091)	142.67[a] (53.976)
Sameclsing	5.414[a] (1.582)	3.424[a] (1.230)
Sameclsed	1.877[b] (0.505)	1.199 (0.392)
Same assignee	11.432[a] (3.459)	
Number of obs.	4803	3891
Log-likelihood	-7.076	-3.493
Pseudo-R2	0.453	0.383

For explanatory notes see the legend of Table 28.1.The baseline category left out is represented by patent pairs not socially connected (disconnected) and spatially separated. Only significant coefficients are reported.
[a] significant at the 1% level; [b] significant at the 5% level.

Table 28.3 generalises these results by looking at the interaction effects between social distance and geographical co-location. Once again, our findings suggest that for low social distances, social networks represent channels of inter-firm knowledge diffusion that are probably more effective than geographical co-location per se. Looking particularly at the second column of table 28.3, which excludes company self-citations, one observes that the probability for patent pairs linked by paths shorter than 2, independently on the spatial location, is several orders of magnitude higher than for patent pairs linked by longer paths or even not connected. At the same time, however, one should not overlook that geography still plays a role in mediating knowledge flows. For unconnected patent pairs the odds of a citation link are still 6–7 times higher for spatially co-located patent pairs than for non co-located pairs.

5. CONCLUSIONS

The application of patent citation analysis to the study of knowledge diffusion has not yet reached firm conclusions. Controversy still flourishes over whether the interpretation of citations as 'paper trails' left by interpersonal knowledge flow is legitimate.

In this paper we have stressed that much of the debate depends on the popularity and the peculiarities of the US patent system, whose limitations may not affect the European system. In particular, many efforts placed on distinguishing between inventors' and examiners' citations are pointless when using EPO data. We have also stressed how the same efforts can indeed be judged excessive for the purpose of using USPTO data, as long as one recognises that knowledge of the technical contents of a patent may travel independently from information about the existence of that patent, or from exact references to the relevant documents.

Social network analysis can be more decisive, for at least three reasons. First and foremost, because it recognises that information may travel from inventor to inventor not only directly, but also indirectly, via complex social chains. Second, because inventors, at least in R&D and patent intensive fields, may well represent a 'community of experts', that is a meaningful unit of analysis. Third, because a methodology has emerged from the recent literature, which allows us to test the influence of social distance on citation probabilities. When applied to EPO data, that methodology confirms that short social chains of inventors are indeed more likely to generate citations than unconnected inventors.

Much more research ought to be done, both to refine the concept of 'social network', when applied to communities of inventors, and to improve the quality of data, both about inventors, applicant companies and citations. However, social network analysis of patent citations looks as one of the most promising research avenues in the field of innovation and diffusion studies.

ACKNOWLEDGEMENTS

Research leading to this paper has been supported by a grant provided by the research directorates of the European Commission to the *ineck* thematic network.

REFERENCES

Agrawal, A.K., Cockburn I.M., McHale J. (2003). *Gone but not forgotten: labor flows, knowledge spillovers, and enduring social capital*. NBER Working Paper 9950.

Akers, N. (1999). The European Patent System: an introduction for patent searchers, *World Patent Information,* 21, 135–163.

Akers, N. (2000). The referencing of prior art documents in European patents and applications, *World Patent Information,* 22, 309–315.

Almeida, P., Kogut, B. (1999). The localization of knowledge and the mobility of engineers in regional networks. *Management Science,* 45, 905–917.

Balconi, M., Breschi, S., Lissoni, F. (2004). Networks of inventors and the role of academia: an exploration of Italian patent data. *Research Policy* (forthcoming).

Breschi, S., Lissoni, F. (2001a). Knowledge spillovers and local innovation systems: A critical survey. *Industrial and Corporate Change*, 10 (4), 975–1005.

Breschi, S., Lissoni, F. (2001b). Localised knowledge spillovers vs. innovative Milieux: knowledge tacitness" reconsidered. *Papers in regional science,* 80 (3), 2001.

Breschi, S., Lissoni, F., Malerba F. (2003). *STI-NET patent and patent citations database. Methodology and preliminary analyses*. Mimeo.

Constant, E.W. (1984). *Communities and hierarchies: Structure in the practice of science and technology*. In R. Laudan (Ed.), The nature of technological knowledge. Dordrecht: D. Reindel Publishing, S. 27–46.

Cowan, R., David, P.A., Foray, D. (2000). The explicit economics of knowledge codification and tacitness. *Industrial and Corporate Change,* 9, 211–253.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature,* 28, 1661–1707.

Hall, B.H., Jaffe, A., Trajtenberg, M. (2000). *Market value and patent citations: A first look*. NBER Working Paper 7441.

Hall, B.H., Jaffe, A., Trajtenberg, M (2001). *The NBER patent citations data file: lessons, insights and methodological tools*. NBER Working Paper 8498. Republished in: Jaffe and Trajtenberg (2002).

Jaffe, A.B., Trajtenberg, M. (2002). *Patents, citations, and innovations: A window on the knowledge economy*. MIT Press.

Jaffe, A.B., Trajtenberg, M., Fogarty, M.S. (2000a). *The meaning of patent citations: Report on the NBER/Case-Western reserve survey of patentees*. NBER Working Paper 7631.

Jaffe, A.B., Trajtenberg, M., Fogarty, M.S. (2000b). Knowledge spillovers and patent citations: Evidence form a survey of inventors. *American Economic Review,* 90 (2), 215–218.

Jaffe, A.B., Trajtenberg, M., Henderson, R. (1993). Geographic localisation of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics,* 108, 577–598.

Karki, M.M.S. (1997). Patent citation analysis: A policy analysis tool. *World Patent Information,* 19, 269–272.

King, G., Zeng, L. (2001). Explaining rare events in international relations. *International Organization,* 55 (3), 693–715.

Kortum, S., Lerner, J. (1998). Stronger protection or technological revolution: what is behind the recent surge in patenting? *Carnegie–Rochester Conference Series on Public Policy,* 48, 247–304.

Laboranti, A. (2004). *Il contributo dell'Università di Pavia alle innovazioni tecnologiche,* Degree dissertation, Università degli studi di Pavia, Faculty of Engineering.

Merton, R.K. (1961). *Singletons and multiples in scientific discovery*. Proceedings of the American Philosophical Society, 105, 470–486.
Merton, R.K. (1977). *The sociology of science: An episodic memoir*. Southern Illinois University Press.
Merton, R.K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *ISIS,* 79, 606–623.
Michel J., Bettels B. (2001). Patent citation analysis: A closer look at the basic input data from patent search reports. *Scientometrics,* 51, 185–201.
Møen, J. (2000). *Is mobility of technical personnel a source of R&D spillovers?* NBER Workin Paper 7834.
Murray, F. (2002). Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. *Research Policy,* 31, 1389–1403.
Newman, M.E.J. (2000). *Who is the best connected scientists?* A study of scientific co-authorship networks. SFI Working Paper 00-12-64, Santa Fe.
Newman, M.E.J. (2001). *The structure of scientific collaboration networks.* Proceedings of the National Academy of Science USA 98, 404–409.
Ogburn, W.F., Thomas, D. (1922). Are inventions inevitable? A note on social evolution. *Political Science Quartly,* 3, 83–98.
Pilkington, A., Dyerson, R., Tissier, O. (2002). The electric vehicle: patent data as indicators of technological development. *World Patent Information,* 24, 5–12.
Sampat, B.N. , Ziedonis, A. (2002). *Cite-Seeing: patent citations and economic value of patents.* Mimeo, http://www.vannevar.gatech.edu/paper.htm.
Sforzi, F. (1997). *I sistemi locali del lavoro in Italia. 1991*, Istat, Argomenti n. 10, Istituto Poligrafico e Zecca dello Stato, Roma.
Singh, J. (2003). *Inventor mobility and social networks as drivers of knowledge diffusion*. Mimeo, Harvard Business School.
Sirilli, G. (1987). Patents and inventors: An empirical study. *Research Policy,* 16, 157–74.
Song, J., Almeida, P., Wu, G. (2003). Learning–by–hiring: when is mobility more likely to facilitate knowledge transfer? *Management Science,* 49, 351–365.
Sorenson, O. (2003). *Social networks, informational complexity and industry concentration*. Mimeo, UCLA.
Sorenson, O., Fleming, L. (2001). *Science and the diffusion of knowledge*. Mimeo, Harvard University.
Stolpe, M. (2002). Determinants of knowledge diffusion as evidenced in patent debates: the case of Liquid Crystal Display technology. *Research Policy,* 31, 1181–1198.
Thompson, P. (2003). *Patent citations and the geography of knowledge spillovers: what do patent examiners know?* Mimeo, Carnegie Mellon University.
Thompson, P., Fox-Kean, M.(2003). Patent citations and the geography of knowledge spillovers: A reassessment. Mimeo, Carnegie Mellon University.
Valente, T. (1990).
Wasserman, S., Faust, C. (1994). *Social network analysis: methods and applications*. Cambridge University Press.
Watts, D.J. (2003). *Six degrees: the science of a connected age.* W.W. Norton & Company
Zucker, L.G., Darby, M.R., Armstrong, J. (1998). Geographically localized knowledge: Spillovers or markets? *Economic Inquiry,* 36, 65–86.

Chapter 29

MEASURING THE INTERNATIONALISATION OF THE GENERATION OF KNOWLEDGE

An Approach Based on Patent Data

Dominique Guellec[1] and Bruno van Pottelsberghe de la Potterie[2]

[1]*European Patent Office. E-mail: dguellec@epo.org*
[2]*Université Libre de Bruxelles (ULB) – Solvay Business School, Belgium*

Abstract: This paper presents three new patent–based indicators of internationalisation of knowledge generation. They measure the extent of international co-operation in research and the international location of research facilities associated with multinational firms — i.e., cross-border ownership. These indicators are based on triadic patent data (patent families applied in the US, Europe and Japan), and on the patents granted by the USPTO. They witness both an increasing trend towards the internationalisation of knowledge generation and large cross-country differences in the extent of internationalisation. The degree of technological internationalisation is higher for small countries and for countries with a low R&D intensity. Two countries are more likely to collaborate if they are close to each other, if they have a similar technological specialization and if they share a common language.

1. INTRODUCTION

The internationalisation of research and development (R&D) and innovative activities is an important component of the ongoing trend towards the globalisation of the economy. As industrial operation itself is increasingly conducted on a global basis, technological activities are also involved in this internationalisation trend, although they are probably somewhat lagging behind external trade or international investment in production facilities. At a high level of generality the internationalisation of technology means that inventions, the people generating these inventions, and the ownership of these inventions tend to cross national borders more frequently.

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 645-662.

The objective of this chapter is to measure the extent to which R&D activities are implemented on an international basis. Amongst the various dimensions of the internationalisation of technology we focus on two of them: cross-border ownership of technology (an invention made in country A is owned by a firm based in country B), and the international generation of knowledge (co-operation between industrial R&D laboratories located in different countries).

In what follows we explore measurement and analytical issues relating to the internationalisation of technology, based on patents data. We rely on the methodology developed by Guellec and van Pottelsberghe (2001) to present three new patent–based indicators of internationalisation of technology. These indicators are calculated at the macro-economic level for 34 countries. They relate to the research activity abroad of domestic multinational firms, to the domestic activity of foreign multinational firms and to international co-operation in research. We are primarily concerned with the following issues: Are these indicators different from other indicators of internationalisation of technology (e.g. R&D abroad of multinational firms) or do they convey consistent messages? Has the internationalisation of technology increased over the past 15 years? Does it affect all countries to a similar extent?

The study has been performed on two data bases over the periods 1980–1982 and 1994–1996. The first one is composed of patents granted by the US Patent and Trademark Office (USPTO). The second one is composed of triadic patents (patents that have been applied in the US, in Europe and in Japan). As our indicators are based on patent data, it is worth mentioning that the measure concerns more the internationalisation of applied research and development than basic research (patents codify inventions with a well defined industrial application).

The next section presents a brief state of the art on the literature attempting to measure the internationalisation of technology and of research activities. Section 3 defines the various concepts of internationalisation. Section 4 describes the country patterns of internationalisation of technology. Section 5 summarizes the empirical results of Guellec and van Pottelsberghe (2001) on the determinants of the extent and direction of internationalisation. Concluding remarks are reported in section 6.

2. STATE OF THE ART

An increasing share of technology is owned by firms from a different country than the one of the inventors (which mainly reflects that companies have research facilities abroad). The importance of this phenomenon is not really new. Comparing US patents granted in 1969–72 and in 1983–86, Cantwell (1989, 1995) reported an increasing share of patents with the owner and inventor located in different countries.

Several reasons explain this phenomenon. First, international mergers and acquisitions often end up with research laboratories located in different countries, even if there is no specific technological strategy underlying the fusion process (Patel, 1995 and Niosi, 1999). Second, some multinational firms (MNEs) might set up research facilities abroad in order to adapt their products to local markets and to provide technological support to local subsidiary – the so called home based exploiting strategy. Third is the home base augmenting strategy that reflects the will to monitor new technology developments occurring in foreign countries; to 'tap' into foreign technology. It is not enough to read technical journals to keep pace with advancing technology, it is also necessary to be part of researcher networks. Fourth is an attempt to develop special technology in which the recipient country has comparative advantage and which complements the firm's core technology. While the first three have been validated empirically, the fourth has not (see Patel and Vega, 1999; Dunning and Wymbs, 1999; and Lichtenberg and van Pottelsberghe, 2001).

The internationalisation of the generation of knowledge has been analysed in the literature through two different and complementary approaches. The first one focuses on the internationalisation of research and development activities, whereas the second one relies mainly on patent data. Major contributions to the measurement of the internationalisation of R&D activities have been performed by Cantwell and Harding (1998) for German companies, Cantwell and Iammarino (1998) for Italian companies and several authors for multinational companies (Cantwell and Janne, 1999; Patel and Pavitt, 1991; Patel and Vega, 1999, Zander, 1999; and Gassman and von Zedtwitz, 1999, 2002).

The country level approach taken in this chapter complements the company level approach taken in most of the literature in this field. The company approach is able to capture aspects such as the industry to which the concerned actors belong, it is unambiguous about the nationality of the patentee (the firm) at one point in time. Hence the country level approach cannot address issues such as how does internationalisation relate to the corporate strategy (Is multinational firms' research abroad related to their core activities or to complementary activities?), which is treated in depth in certain firm level studies (Zander, 1995, or Cantwell and Jane, 1999). On the other hand, the country level approach is exhaustive, as all patents are treated, whoever the patentee, instead of a selection of large companies. It further allows more countries to be covered and to give for each country a more complete picture of internationalisation. Moreover, the boundaries of countries are generally stable over time, contrary to those of firms, which facilitates time consistency in the treatment of the nationality of patentees.

3. MEASUREMENT ISSUES

Patents are increasingly recognized as a rich source of information about technological performance. Griliches (1990) and Archibugi and Pianta (1992) provide a good summary of the strengths and weaknesses of patents as indicators of innovative performance and technological specialisation. Amongst the information available from patent files are the inventor and the assignee's countries of residence. When statistically elaborated this information allows us to map some aspects of the internationalisation of technology. The advantages of patents in this area are their broad availability (available for all countries in the world), international comparability (when a few sources of bias are dealt with), and possibilities of being matched with other types of data (firm level, country, or industry level). Their major drawback is a difficulty in interpreting, in some cases, the meaning of the indicators — i.e., which underlying activity is actually reflected in the patent. Dernis, Guellec, and van Pottelsberghe (2001) present an in-depth review of measurement issues of patent indicators.

Patents that are of interest for measuring internationalisation of technology are those with several applicants from different countries, or several inventors from different countries, or an applicant and inventor from different countries. Mapping these populations of patents and comparing them to other patents is the purpose of the indicators which are presented in the next paragraphs.

Cross-border ownership of inventions happens when at least one inventor and the applicant reside in different countries. It is deemed to reflect the location of R&D activities of multinational firms. For most patents (a share usually estimated to be higher than 90%), the applicant is an institution (a firm, a university, a public laboratory). The inventor is always an individual, usually a researcher employed by the applicant. Most often the address of the inventor is the address of the laboratory he/she works in. Then when the inventor and the applicant of a patent do not reside in the same country, this reflects in a huge majority of cases that the patent protects an invention performed in a research facility abroad of the headquarter of a multinational firm. Two indicators have been calculated based on this data that mirror each other:

- $\mathbf{I_D A_F}$ is the share for a given country of patents with a domestic inventor and a foreign applicant in the country's total domestic inventions. It reflects the extent to which foreign firms control (own) domestic inventions. Algebraically: P_{ij}^{IA} is the number of patents invented by the residents of country i and at least partly owned by the residents of country j. $P_{i.}^{IA} = \sum_{j \neq i} P_{ij}^{IA}$ is the total number of patents invented by the residents of country i and controlled by foreign residents. $I_D A_{F\,i} = P_{i.}^{IA} / PI_i$ is the share of patents controlled by foreign residents in the total number of patents invented by residents (*PI*).

Patents can also be counted fractionally: when a patent has several inventors residing in different countries, it is 'shared' between these countries, each of which is attributed a fraction which corresponds to its share in the number of inventors. For instance, if a patent has two inventors from country A and three inventors from country B, then country A's number will be 0.4 and country B's will be 0.6. In the present chapter we use the total number of patents invented in a country.

- **$A_D I_F$** is the share for a given country of patents with a foreign inventor and a domestic applicant in the country's total domestic applications. It reflects the extent to which domestic firms control foreign inventions. Algebraically: P_{ij}^{AI} is the number of patents owned by the residents of country i and at least partly invented by the residents of country j. $P_{i.}^{AI} = \sum_{j \neq i} P_{ij}^{AI}$ is the total number of patents controlled by the residents of country i and invented by foreign residents. $A_D I_{F\,i} = P_{i.}^{AI} / PA_i$ is the share of patents invented by foreign residents in the total number of patents controlled by residents (*PA*).

International collaboration in science and technology takes place when a patent has several inventors residing in different countries. This kind of international collaboration between researchers can take place either within a multinational corporation (research facilities in several countries), or through a research joint venture between several firms. It is measured by the following indicator:

- **$I_D I_F$** is the share for a given country of patents with a foreign resident as co-inventor in the population of patents with a domestic inventor. Algebraically: P_{ij}^{II} is the number of patents co-invented by residents of country i and residents of country j. $P_{i.}^{II} = \sum_{j \neq i} P_{ij}^{II}$ is the total number of patents invented by the residents of country i in collaboration with foreign researchers. $I_D I_{F\,i} = P_{i.}^{II} / PI_i$ is the share of patents resulting from international research co-operation in the total number of patents invented by residents of a given country (*PI*).

Table 29.1 presents some illustrative examples of patents concerned with the three indicators. The four patents listed have been granted by the USPTO. They all represent a situation of cross-border ownership. The patents owned by Colgate Palmolive (United States) and Alcatel Alsthom CGE (France) witness a cross-border ownership of invention made by Belgian and German inventors, respectively. One example of both international co-operation between inventors and cross-border ownership is provided by the Microsoft patent, which was invented by researchers with residence in different countries (Franco–US collaboration). The fourth patent illustrates the three types of internationalisation altogether. It is co-owned by a US Hospital and a Belgian University (i.e., an international co-application) and has been invented by two researchers with a Belgian residence and one researcher with residence in the US.

Table 29.1. Examples of cross-border ownership and international co-operation

Appl. No.	*Applicant*		*Inventor*	I_DA_F A_DI_F	I_DI_F
	Name	Resid.	Resid.		
472,807	Microsoft Corporation	U.S.	France (2) U.S. (1)	X	X
859,431	Colgate Palmolive Comp.	U.S.	Belgium (2)	X	
828,191	Alcatel Alsthom CGE	France	Germany (2)	X	
463,418	General Hospital Corp. Rijksuniversiteit	U.S. Belgium	Belgium (2) U.S. (1)	X	X

Source: see Guellec and van Pottelsberghe (2001), USPTO.

These various indicators are not independent of each other. For instance, all patents with co-inventors residing in different countries will also have, for at least one of the concerned countries, a foreign applicant. As a consequence any patent counted in I_DI_F for a country will also be counted either in I_DA_F or in A_DI_F for this country. The share of co-inventions in the population of foreign owned patents allows to measure the extent to which cross border ownership favours international circulation of knowledge. Another feature is that world wide I_DA_F is equal to A_DI_F. This fact does not preclude large differences between the two indicators at the country level.

An indicator similar to A_DI_F has been proposed by the University of Reading, relying on patent count data obtained from the Science Policy Research Unit (SPRU; see especially Patel and Pavitt, 1991, 2000; Dunning, 1994; and Dunning and Wymbs, 1999). This indicator has been computed with data about the patents registered in the US by 727 of the world's largest firms. It reflects the share of US patents filed by these firms attributable to research in foreign locations (outside the home country of the parent company). Contrary to SPRU, we use all patents (including those filed by small firms, large domestic firms, public institutions, universities, etc.) which means that the degree of internationalisation of a country is represented more completely. However, identifying the owner firm as SPRU does — and we do not — allows the actual owner of the patent to be identified, beyond the direct owner: hence foreign affiliates of multinational firms can be identified and treated as such, which is not the case in our approach. In any case, it is worth mentioning that cross-country and cross-industry differences observed with the SPRU indicator are similar to those observed with A_DI_F.

The indicators proposed here should be considered as lower bounds indicators of internationalisation of technology. A major factor of under-estimation of actual internationalisation in our data is related to the increasing tendency, since the early 1980s at least, of cross-border mergers and acquisitions (M&A). As underlined by Dunning (1994), an important *raison d'être* for the growing share of these M&As is to acquire the innovative capacity of the targeted firms. Patent databases do not register such changes in ownership of patents.

In what follows we use a different data source from Guellec and van Pottelsberghe (2001), who relied mainly on patent applications at EPO. The indicators presented in the next section are based on patents granted by the USPTO and on triadic patents. The latter are patent families which are applied simultaneously at least at the USPTO, the EPO and the Japanese Patent Office (JPO). The reason underlying the use of triadic patents is that they reflect inventions with a higher potential economic value (otherwise they would not be applied around the globe) than the patents that are only applied in one regional patent office.

4. GLOBAL TRENDS AND COUNTRY PATTERNS

Table 29.2 presents the three indicators for most triadic patents and the patents granted by the USPTO. It clearly appears that the internationalisation process takes more the form of cross-border ownership than collaboration between inventors of different countries. From 1980 to 1995, the degree of international R&D collaboration (I_DI_F, the share of triadic patents involving at least two inventors from different countries) has more than doubled, from 2.5% in 1980 to 5.8% in 1995 (for patents granted by the USPTO the shares of co-inventions were 1.2% and 3.3%, respectively). In other words, 6 out of 100 patented inventions are the fruit of international research collaboration.

The share of cross-border ownership has been quite stable in the 1980s, fluctuating around 10% of triadic patents. From 1990 onwards cross-border ownership of inventions has grown steadily to reach a share of 15%. More than one out of 10 triadic patents is subject to cross-border ownership. This is similar to the share of foreign affiliates of multinational firms in R&D expenditure, which averages about 15% OECD-wide.

Table 29.2. The internationalisation of technology: global trends, %

	$I_D I_F$		$I_D A_F / A_D I_F$	
Type of patent	1980	1995	1980	1995
Triadic	2.5	5.8	10.0	15.0
USPTO	1.2	3.3	5.2	8.5

Source: OECD, own calculations, by priority year.

Beyond these global trends countries display various patterns of internationalisation (the cross-country average share is higher than the worldwide share since in the latter case there is some clearing: each patent co-invented in several countries is counted only once, whereas in the former case it is counted *in each country*). Table 29.3 shows the evolution from the early eighties to the mid-nineties of the three indicators based on triadic patent data, for 35 countries. Table 29.4 compares the three indicators in the mid-nineties computed either with patents granted by the USPTO or with triadic patents. Five observations can be drawn from these tables.

First, there is a striking heterogeneity across countries. For instance, very few inventions made in Japan are controlled by foreign firms, invented in co-operation with foreign researchers, or controlled jointly with a foreign applicant. Similarly, only few inventions controlled by Japanese residents are invented abroad. With the exception of Japan, it is clear that a significant proportion of patents are subject to cross-border ownership and international collaboration. Amongst the largest patenting countries the United Kingdom is characterised by a relatively high degree of internationalisation of its technology, with ratios ranging between 20 and 50 per cent. In the early eighties (mid-nineties) 6% (23%) of its triadic patents were invented with foreign researchers; 29% (46%) of the triadic patents invented domestically were controlled by a foreign firm; and 7% (21%) of the inventions controlled by firms based in the United Kingdom were invented abroad. Smaller countries (such as Belgium, Austria and Ireland) and/or less developed countries (Turkey, Mexico, Poland) are also highly internationalised.

Table 29.3. Three indicators of internationalisation of technology, triadic patents, %

Country	I_DI_F 80–82	I_DI_F 94–96	I_DA_F 80–82	I_DA_F 94–96	A_DI_F 80–82	A_DI_F 94–96
Europe						
Austria	.16	.31	.37	.47	.08	.20
Belgium	.23	.32	.43	.48	.10	.23
Czech Republic	-	.50	-	.67	-	.18
Denmark	.14	.24	.19	.27	.08	.22
European Union	.03	.09	.16	.21	.03	.08
Finland	.05	.11	.07	.08	.07	.18
France	.05	.14	.12	.22	.05	.12
Germany	.06	.12	.12	.17	.05	.10
Hungary	.01	.42	.01	.47	.02	.16
Ireland	.45	.51	.45	.63	.40	.59
Italy	.07	.14	.17	.26	.03	.07
Luxembourg	.19	.65	.28	.78	.37	.83
Netherlands	.08	.22	.56	.67	.21	.44
Norway	.12	.23	.21	.25	.10	.24
Spain	.16	.34	.25	.42	.12	.10
Sweden	.05	.16	.14	.19	.06	.18
Switzerland	.15	.30	.42	.35	.29	.39
United Kingdom	.06	.23	.29	.46	.07	.21
North America						
Canada	.20	.29	.28	.36	.17	.29
Mexico	-	.65	-	.79	-	.42
United States	.03	.09	.03	.07	.17	.19
South America						
Argentina	-	.35	-	.48	-	.21
Brazil	-	.40	-	.54	-	.13
Asia						
China	-	.44	-	.51	-	.26
Hong Kong - China	-	.42	-	.62	-	.36
India	-	.73	-	.83	-	.18
Japan	.01	.03	.02	.04	.01	.03
Korea	-	.07	-	.09	-	.07
Singapore	-	.45	-	.86	-	.31
Taiwan	-	.59	-	.57	-	.51
Others						
Israel	.12	.28	.28	.41	.10	.15
New Zealand	.00	.31	.00	.39	.03	.20
Russian Federation	-	.52	-	.80	-	.31
Australia	.12	.20	.18	.34	.05	.16
South Africa	.09	.33	.24	.61	.06	.21

Source: OECD, own calculations, by priority year.

The bottom lines of Table 29.4 provide the median of the three indicators across countries. It clearly shows that the countries with less than 100 triadic patents invented domestically are much more involved in international collaboration and see a larger share of their invented patents controlled by foreign companies. It is already well known that small economies are more internationalised than large ones in terms of trade as well as foreign direct investment. It turns out that this applies to technology as well. The ‘technological size’ of a country seems to have a close relationship with its degree of internationalisation. Concerning the share of research co-operation, this may partly be explained by each researcher from a small country having fewer local colleagues in the field and must therefore look abroad for collaboration.

Second, the three indicators show that there has been a significant increase, for all countries, in the level of internationalisation between the early eighties and the mid-nineties.

Third, it seems that R&D intensive countries are much less internationalised than the other countries. Table 29.4 shows that the median indicator for the eight most R&D intensive countries ranges from 10 to 15 percent (whatever the type of indicator and the data source). For all countries the median indicator ranges from 14 to 46 per cent.

Table 29.4. Three indicators of internationalisation of technology, 1994–1996, %

	I_DI_F		I_DA_F		A_DI_F	
Selected countries	*USPTO*	*TRIAD*	*USPTO*	*TRIAD*	*USPTO*	*TRIAD*
China	.38	.44	.54	.51	.25	.26
European Union	.10	.09	.21	.21	.08	.08
France	.13	.14	.21	.22	.11	.12
Germany	.11	.12	.16	.17	.09	.10
Japan	.02	.03	.03	.04	.03	.03
United States	.04	.09	.04	.07	.08	.19
Median						
All	.23	.31	.34	.46	.14	.20
Less than 100 triadic patents	.31	.44	.43	.62	.17	.21
Between 100 and 1000 triadic patents	.23	.26	.31	.37	.12	.19
More than 1000 triadic patent	.14	.15	.22	.24	.13	.18
8 RD intensive (*)	.12	.11	.13	.13	.10	.15

Source: OECD, own calculations, by priority years. * indicates the eight most R&D intensive countries: Finland, France, Germany, Japan, Korea, Sweden, Switzerland and the USA. Median figures are based on the 35 countries listed in Table 29.3.

Fourth, in most countries the share of domestic inventions owned by foreign firms (I_DA_F) is substantially higher than the share of foreign inventions in total domestic applications (A_DI_F). The reverse is true for only five countries. This is owing to a concentration of ownership of cross-border patents in the hands of a few countries. Actually, four countries are the largest owners of patents covering foreign inventions: the United States (although, because of its size, the *share* of foreign inventions is just under the median level: but the *level* is high), Switzerland, Finland, and Sweden. These are also countries with well known, strong multinational firms.

Fifth, there seems to be a systematic bias between the two sources of data. Patents filed by residents of non-European countries are more internationalised with triadic patents than with USPTO patents. For instance, about 4 (24) per cent of the patents invented by residents of the United States (Canada) and granted by the USPTO are controlled by foreign companies, against 7 (36) per cent with triadic patents. The main reason is that the higher 'proximity' of the USPTO to the United States and Canada leads them to patent there their purely domestic inventions to a larger extent than in other patent offices ('home advantage').

Cross-industry differences are also reported in Guellec and van Pottelsberghe (2001). The variance is significantly lower across industries than across countries, suggesting that the internationalisation of technology is more related to country peculiarities than technological ones. Four sectors are nevertheless highly internationalised: Chemicals, Oil refining, Drugs, and Food and Beverages. Shipbuilding and Aerospace, two sectors generally considered to be subject to special government attention are the least internationalised. These differences across industries confirm the findings of Dunning and Wymbs (1999). Their survey of the world's largest multinationals shows that pharmaceutical firms obtain more of their competitive advantage from foreign sources than other sectors, whereas firms from the aerospace sector rely the most on domestic sources.

How similar to each other are the three patent–based indicators of internationalisation of technology? And are they similar to other indicators of internationalisation? In other words, do the different dimensions of internationalisation go together, so that certain countries are more opened than others in all respects – or, conversely, are the different dimensions of internationalisation substitutes for each other, allowing different patterns of internationalisation to take place? Table 29.5 reports cross-country correlation between these indicators. All pairs of indicators related to the internationalisation of inventions are significantly correlated with each other. The highest correlation (0.83) is for the pair I_DI_F - I_DA_F: the higher a country's share of domestic inventions controlled by foreign companies, the more it collaborates with foreign countries.

The three indicators (especially I_DA_F) are highly correlated with the share of production by affiliates of foreign owned firms in total domestic production. Only I_DA_F and I_DI_F have a significant correlation with the share of foreign affiliates in total domestic R&D expenditure. That is, international collaboration and foreign ownership of domestic inventions are closely related to the innovative and production activities of foreign affiliates.

All the patent related indicators are positively correlated with openness to external trade (imports and exports relative to GDP). The only indicators that do not seem to provide similar pictures of internationalisation are FDI (both inward and outward). This is probably because statistics on FDI mainly report net flows of FDI and not the stock of FDI.

Table 29.5. Cross-country correlation between various indicators of internationalisation

Country	*nobs*	I_DA_F	A_DI_F	I_DI_F
I_DA_F	20		0.47	0.83*
A_DI_F	20	0.47		0.55
I_DI_F	20	0.83*	0.55	
Activities of foreign affiliates				
Share of foreign output (SHFP)	12	0.86*	0.83*	0.83*
Share of foreign R&D (SHFR)	11	0.90*	0.56	0.89*
Foreign direct investments				
Share of inward FDI (SHINF)	20	0.25	0.13	0.32
Share of outward FDI (SHOUF)	20	0.01	0.41	0.07
International trade				
Share of imports in GDP (MGDP)	20	0.76*	0.69*	0.72*
Share of exports in GDP (XGDP)	20	0.69*	0.75*	0.66*

Correlation across 20 OECD countries with at least 200 EPO patents invented over the period 1993–95. nobs indicates the number of available observations for each variable; I_DA_F the share of domestic inventions with foreign applicants; A_DI_F the share of domestic applications with foreign inventors; I_DI_F the share of domestic inventions with at least one foreign inventor; SHFP the share of domestic output produced by foreign firms; SHFR the share of domestic R&D in foreign firms; SHINF the share of inward FDI in gross fixed capital formation; SHOUF the share of outward FDI in gross fixed capital formation; MGDP the share of imports in GDP; and XGDP the share of exports in GDP; * indicates the coefficients that are significant at a 5% probability threshold.
Source: Guellec and van Pottelsberghe (2001).

5. THE INTENSITY AND THE DIRECTION

What are the common determinants of the openness of countries to foreign technology? The few studies on the determinants of internationalisation of technology have essentially focused on the share of foreign R&D, at the firm level (see the survey by Granstrand *et al.,* 1992). The major determinants are the age of the firm, its size, its stage of corporate development and its international pattern of manufacturing. Guellec and van

Pottelsberghe focused on two main factors for each country: the relative level of technological endowment, proxied by the research intensity (GERD/GDP ratio), and the size of the country (GDP). Both variables were taken in logarithmic form, in order to capture non linear effects:

$$X_D X_{F\,i} = c + \beta \log\left(GDP_i\right) + \alpha \log\left(IRD_i\right) + e_i \quad , \tag{1}$$

where X stands for A or I (i.e., applicant or inventor). The dependent variable is the indicator of internationalisation, c is an intercept, e is the error term, α and β are the parameters to be estimated, i is the country index. The econometric estimates for each of the three indicators (with EPO and USPTO patents) led to the following observations.

First, the various types of internationalisation do not seem to respond to the same determinants. However, in all cases the effect of GDP is negative (although not always significant), showing that smaller countries are more internationalised than larger ones *ceteris paribus*. R&D intensity has a negative impact on I_DI_F and I_DA_F, and a positive one on A_DI_F. The extent of collaboration with foreign researchers (I_DI_F) is very well explained by the model. The higher (lower) the R&D intensity of a country the less (more) its researchers enter into collaboration with foreign colleagues. Countries with a low R&D intensity and of a small size rely more on external co-operation owing to their own, weaker, capabilities, therefore relying more on knowledge flows from abroad.

I_DA_F also decreases with R&D intensity: the higher (lower) the relative R&D spending of a country the lower (higher) is the share of its residents' invention that is controlled by multinational firms. In other words, national control over domestic inventions increases with domestic inventive efforts. There is only a small (statistically insignificant) negative effect of GDP. A_DI_F is less well explained by the model. The size of a country is a negative and significant determinant of the extent to which it controls foreign inventions, whereas its technological intensity is a positive but not significant determinant. The smaller a country is the higher is the share of inventions it owns that are invented abroad. In a nutshell, the more a country is intensive in research the less its own inventions are controlled by foreign firms and the less it enters into international research co-operation. The larger the country the lower is its share of patent applications that have been invented abroad and the lower is its propensity to enter into international co-operation.

The asymmetry between I_DA_F and A_DI_F tends to confirm the idea that the firms based in leading edge countries exploit their technological advantage more through foreign acquisition (cross-border ownership), through A_DI_F. The negative sign of the relationship between I_DA_F and R&D intensity does not support the argument that leading edge countries are being 'techno-sourced', at least not through foreign ownership of their own invention facilities.

Guellec and van Pottelsberghe (2001) also analysed the geographical distribution of the internationalisation of technology: with which partners do each country tend to co-operate more, and less? For instance, regarding the patents invented by residents and owned by foreign applicants (I_DA_F), the share of US residents is higher in non-European countries, except Ireland and the United Kingdom. The share of Japan is higher in Korea, Australia, and the United States (but this is partly a statistical artefact since the United States is not partner to… the United States, which tends to inflate the share of all countries in the United States compared with their share in other countries). For each European country (except the United Kingdom, Ireland, and Luxembourg), the highest share of foreign owned patents goes to other European countries. Similar patterns appear for A_DI_F and I_DI_F.

In order to control for the size effect (the United States is the largest partner to all countries, simply because it is the biggest patenting country of all) an index of 'revealed geographical distribution' (RGD) can be computed for each of the three indicators of internationalisation of technology. This index is similar, in spirit, to the 'revealed comparative advantages' that international economists are familiar with. Basically, it is country *j*'s share in country *i*'s foreign relationships relative to its share in OECD. For instance, it is the share of German residents in French inventors' patents owned by foreigners, divided by German residents' share in total OECD foreign owned patents. Algebraically, we have:

- The RGD of foreign ownership of domestic inventions is equal to the share of country *j* in country *i*'s patents owned by foreign residents divided by the share of country *j* in the world wide patents subject to cross-border ownership: $RGD\_IA_{ij} = \left[P_{ij}^{IA} / P_{i.}^{IA}\right] * \left[P_{.j}^{IA} / P_{..}^{IA}\right]$.
- The RGD of domestic ownership of foreign inventions is equal to the share of country *j* in country *i*'s patents owned by foreign resident divided by the share of country *j* in the world wide patents subject to cross-border ownership: $RGD\_AI_{ij} = \left[P_{ij}^{AI} / P_{i.}^{AI}\right] * \left[P_{.j}^{AI} / P_{..}^{AI}\right]$. By construction, *RGD_AI* is the transposed matrix *RGD_IA*.
- The RGD of international co-inventions is equal to the share of country *j* in a country *i*'s patents co-invented with foreign residents divided by the share of country *j* in the world wide patents subject to international collaboration:

$$RGD\_II_{ij} = \left[P_{ij}^{II} / (P_{i.}^{II} - P_{ij}^{II})\right] * \left[(P_{.j}^{II} - P_{ij}^{II}) \Big/ (\sum_i P_{i.}^{II} - P_{i.}^{II} - P_{.j}^{II}) \right]$$

These indicators were calculated for each pair of OECD countries with patent applications at the EPO, resulting in three 29*29 matrices. We identified and tested five factors that may explain the revealed geographical distribution of the different indicators of internationalisation of technology. The first one is the technological proximity (*TP*) between pairs of countries. To measure the proximity of countries *i* and *j* we use the uncentered correlation of the two countries' distribution vectors of patents across 30 technological classes in 1992–95 (*Fi* and *Fj*), as follows:

$$TP_{ij} = F_i F_j' \, / \, [(F_i F_j') \, (F_{ji})]^{1/2} \, .$$

This indicator is equal to one for the pairs of countries whose technological specialisations are identical, it is equal to zero for pairs of countries whose vectors are orthogonal, and it is bounded between 0 and 1 for all other pairs of countries. It is similar to Jaffe's (1986) indicator of technological proximity between US firms.

The second explanatory factor is the geographical distance (*DGD*), proxied by a dummy variable taking the value one if countries *i* and *j* have a common border, and 0 otherwise. Then come two particular dummy variables reflecting that countries *i* and *j* are a member or not of the European Union (DEU, included 12 countries for the two sub-periods; for the European Union member countries, the proximity dummy variable accounts for 25% of the European Union dummy; about 25% of the pairs of EU Member countries share a common border), whether they are both Nordic countries (*DNORD*) and whether they share a common language (*DLANG*), be it English, Spanish, or German. These dummies aim at testing whether the common membership to the European Union or common languages (and hence cultural and historical similarities) increase the propensity of firms from two different countries to collaborate with each other.

The estimates provided fairly good results that led to the following observations. First, technological proximity matters. The closer two countries are in their technological specialisation, the more they co-operate in research and hold patents invented by researchers of the other country. Second, the geographical distance has a significant effect, for the two types of internationalisation of technology. Countries with a common border enter into cross-border ownership and co-operate more. In the age of globalisation geography still matters (along with history, which is more shared between closer countries). The common history and relatively weak cultural differences that characterise the Nordic countries is probably the main explanation of the positive and significant parameter associated with *DNORD*. Third, countries which share a common language co-operate more with each other and have a higher propensity of entering into cross-border ownership. Finally, pairs of countries which are both members of the European Union have slightly more cross-border ownership of patents, but not more research co-operation.

This positive relationship between cross-border ownership and European Union membership appeared between the mid-1980s and the mid-1990s. One interpretation would be that the policies fostering the European integration have stimulated a process of industrial and financial concentration (through M&As) between European firms. Such cross-border consolidation translates into more cross-border ownership of patents, as the new firm owns research facilities in different countries. European countries co-operate more with each other than with non-European ones, but not more than their language similarity and their geographical and technological proximity would predict. Hence industrial and financial concentration has not yet resulted in closer research links: there is not yet real integration of European countries in the field of business R&D.

6. CONCLUDING REMARKS

This chapter has presented and analysed three indicators of internationalisation of technology derived from the information available in patent data. These indicators have been first developed by Guellec and van Pottelsberghe (2001) for OECD countries and were mainly based on patent applications at the EPO. In this chapter the indicators were derived from triadic patent families (considered to be of a high economic value) and from the patent granted by the USPTO for 35 countries involved in significant patenting activity. Triadic patents, generally considered to be associated with a higher economic value, consistently show higher levels of internationalisation than the patents granted by the USPTO.

In accordance with the existing literature the indicators witness an increasing trend towards the internationalisation of technology for all countries. However, there are large differences in the extent of internationalisation across countries. Internationalisation of a country’s technological activities decreases with its size and with its R&D intensity. Researchers in larger countries more easily find colleagues for partnering in their own country, and countries with higher technological level do not need as much as others co-operation with foreign researchers since their own knowledge base is large. This partly explains the relative insulation of Japan for instance. The spoken language may also be part of the explanation, as suggested by the results concerning the geographical patterns of international co-operation. Another insight is that these indicators are closely correlated with other indicators of internationalisation of technology, such as the share of output produced by foreign firms, or the share of R&D performed by foreign firms.

Who co-operates with whom is largely explained by geographical proximity and technological proximity (similar specialisation) of the partnering countries. In addition, sharing a common language fosters bilateral links in technology. Pairs of countries which are both members of the European Union have slightly more cross-border ownership than the average, but not more research co-operation than is implied by their geographical and technological proximity. Narula (2003) confirms these results with its own analysis of European firms' R&D cooperation patterns, which is higher with US companies than with other European countries.

This chapter shows that patent–based indicators of cross-border ownership and of international collaboration allow us to obtain a relevant, although imperfect, measure of the internationalisation of technology, both over time and across countries. The major drawback of our indicators is that they do not take into account the impact of international mergers and acquisitions and therefore underestimate the real level of internationalisation of the generation of knowledge.

REFERENCES

Archibugi, D., Howells, J., Michie, J. (1999). *Innovation policy in a global economy*. Cambridge University Press.

Archibugi, D., Iammarino, S. (1999). *The policy implications of the globalisation of innovation*, chapter 12. In Archibugi et al. (Ed.), (pp. 242–271).

Archibugi, D., Pianta, M. (1992). *The technological specialisation of advanced countries,* A report to the EEC on international science and technology activities. Boston: Kluwer.

Cantwell, J. (1989). *Technological innovation and multinational corporations*. New York: Blackwell.

Cantwell, J. (1995). The globalisation of technology: what remains of the product life cycle model. *Cambridge Journal of Economics*, 19 (1), 155–174.

Cantwell, J., Harding, R. (1998). The internationalisation of German companies R&D. *National Institute Economic Review*, 163, 99–124.

Cantwell, J., Iammarino, S. (1998). MNCs, technological innovation and Regional Systems in the EU: Some evidence in the Italian case. *International Journal of the Economics of Business,* 5 (3), 383–408.

Cantwell, J., Janne, O. (1999). Technological globalisation and innovative centres: the role of corporate technological leadership and locational hierarchy. *Research Policy,* 28, 119–144.

Dernis, H., Guellec D., van Pottelsberghe, B. (2001). Using patent counts for cross-country comparisons of technology output. *STI Review*, 27, 129–146.

Dunning, J.H. (1994). Multinational enterprises and the globalization of innovatory capacity. *Research Policy*, 23, 67–88.

Dunning, J.H., Wymbs, C. (1999). *The geographical sourcing of technology–based assets by multinational enterprises*, chapter 10. In Archibugi et al. (Ed.), 185–224.

Gassman, O., von Zedtwitz, M. (1999). New concepts and trends in international R&D organisation. *Research Policy*, 28, 231–250.

Gassman, O., von Zedtwitz, M. (2002). Market versus technology drive in R&D internationalisation: four different patterns of managing research and development. *Research Policy,* 31, 569–588.

Granstrand, O., Häkanson, L., Sjölander, S. (1993). Internationalization of R&D. A survey of some recent research, *Research Policy*, 22, 413–430.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature*, 28, 1661–1707.

Guellec, D., van Pottelsberghe, B. (2001). The internationalisation of technology analysed with patent data. *Research Policy*, 30 (8), 1256–1266.

Jaffe, A.B. (1986). Technological opportunity and spillovers of R&D: Evidence from firm's patent, profits and market value. *The American Economic Review*, 76 (5), 984–1001.

Lichtenberg, F., van Pottelsberghe, B. (2001). Does foreign direct investment transfer technology across borders? *The Review of Economics and Statistics*, 83 (3), 490–497.

Narula, R. (2003). *Technology and globalisation: Interdependence, innovation systems and industrial policy*. Cambridge University Press.

Niosi, J. (1999). The internationalization of industrial R&D: From technology transfer to the learning organisation. *Research Policy*, 28, 107–117.

Patel, P. (1995). Localised production of technology for global markets. *Cambridge Journal of Economics* 19 (1), 141–154.

Patel, P., Pavitt, K. (1991). Large firms in the production of the world's technology: An important case of non–globalization. *Journal of International Business Studies*, 22 (1), 1–22.

Patel P., Pavitt, K. (2000). National systems of innovation under strains: The internationalization of corporate R&D. *Miméo*, SPRU, University of Sussex.

Patel, P., Vega, M. (1999). Patterns of internationalization of corporate technology: location vs. home country advantages. *Research Policy*, 28, 145–155.

Zander, I. (1995). *Technological diversification in the multinational corporation — Historical evolution and future prospect*. In: Schiatarella R. (Ed.), New challenges for Europe and international business, Cofindustria, Rome.

Zander, I. (1999). How do you mean "global"? An empirical investigation of innovation networks in the multinational corporation. *Research Policy*, 28, 195–213.

Chapter 30

PATENTS AND PUBLICATIONS

The Lexical Connection

Elise Bassecoulard[1] and Michel Zitt[1,2]

[1]*INRA, Lereco, Institut National de la Recherche Agronomique, Nantes, France.*
E-mail: bassecou@nantes.inra.fr
[2] *OST, Observatoire des Sciences et des techniques, Paris, France*

Abstract: The quantitative appraisal, partly through bibliometrics, of science–technology connections has made great progress in the last decade. We investigate in this chapter the lexical linkage between articles and patents, an alternative method to the systematic exploitation of the citations of patents to scientific papers. We explore in particular the ability to establish correspondence tables between patent classification and scientific categories. After a reminder of the methodological background (S&T linkages, lexical methods, statistical measures) we report an exploratory study based on a subset of the Chemical Abstracts database (CA) that covers both articles and patents by a very precise indexing system. Connection measures have been established, first on controlled vocabulary, and secondly on some natural language fields. The comparison shows some robustness of the lexical approach, with clear limitations at the micro level: topic sharing between a particular article and a particular patent cannot be interpreted in the general case as the sharing of a research question. At the macro level, for example IPC sub-classes and ISI subject categories, the lexical approach is an appealing technique, complementary to usual citation based analysis built on very sparse matrices, because informetric performances of lexical methods can be tuned in a large scope of precision–recall features. The extension to databases specific either to articles or patents requires language processing which can be alleviated if macro level correspondence is solely sought.

1. INTRODUCTION

The science–technology relation and the university–industry relations are crucial issues for understanding the knowledge society, address science policy questions, and provide tools for S&T watch. The traditional analysis

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 665-694.

relies on the separation between science, a public good produced by a self-organised community in universities on the one hand, and technology, an IPR protected good in charge of the industry on the other hand, with a descending flow of knowledge. The emerging knowledge based economy needs more complex descriptions where the distinctions science/technology and universities/industry are neither absolute nor co-extensive. The tight fabric of relations between actors involved in STI has been described within models or metaphors such as Mode 2 (Gibbons et al., 1994) and Triple Helix (Etzkowitz and Leydesdorff, 1997) with examples in biotechnology, bioinformatics, and electronics. The traditional model of free science could be challenged by societal pressure and IPR extension (Dasgupta and David, 1994; David and Foray, 1995) and the view of science as a public good is even questioned (Callon, 1994).

Bibliometrics is one family of methods, amongst many others (surveys, expert groups, etc.) able to provide measures of the linkage between science and technology at micro and macro scale. The bibliometric approach can process scientific publications on the one hand, patents on the other hand, considered as support of information amenable to statistical treatment. S&T indicators based on bibliometrics assume: (a) that publications are a good representation of the contents of science (this hypothesis relies on the sociologists' and historians' works); (b) that patents (or kindred IPR forms) collect a large part of technological information, even though the patent combines several functions. These hypotheses have been largely discussed in the literature: both forms are proxies, they are not fully representative, their value distribution is very skewed, etc. It remains that publication archives on the one hand, patent systems on the other hand, represent a great memory of theoretical and practical knowledge and know-how, accessible, moreover, with a high degree of codification that allows quantitative analysis. Transmission of tacit knowledge via face-to-face interactions, sharing scientific instruments etc, needs other methods of investigation (see the chapter by Tijssen in this handbook).

Publications and patents as information supports have many analogous features (author/inventor, institution/assignee, bibliographic referencing/ patent system referencing, bibliometric classification/official classification, abstract, full text, references to scientific literature/references to patent or non-patent literature, etc.). Some databases store both forms of documents. Scholars have established further informetric analogies (for example, statistical distributions of productivity, of citation/value), which justify extending bibliometric techniques to patents, including the core of bibliometrics, citation analysis, with respect to fundamental differences in status and interpretation (Pavitt, 1985; Narin, 1994; Grupp, 1998).

In this logic, informetric study of the linkage publication–patent is only one particular case of bibliometric data analysis and mapping, and the various entry points to investigate this linkage are the classic tetralogy: citation networks; authors/institutions network; classification schemes; lexical network – and their combination. Naturally, this particular case of bibliometrics should particularly pay attention to the discrepancies between the two entities.

In this chapter we shall focus on one family of method, the study of the lexical connection. Section 2 is devoted to the general context: stakes of the measures of linkage, recall of the various bibliometric ways of addressing this question, especially the now classic approaches by citations and co-activity. Section 3 details the issues of lexical description of publications and patents, and of methods applicable to dealing with these descriptions. Section 4 reports a particular experience[1] intended to assess the feasibility of a lexical approach to issue of the concordance of scientific and technological nomenclatures. Section 5 is devoted to a discussion and conclusion.

2. BACKGROUND

2.1 Applications

The measure of the linkage between patents and publications is appealing in many respects. It helps the understanding of intensity, orientation, and of sources of the science–technology relation. As noted above, the distinctions university/industry and publications/patents are less co-extensive than ever, and kindred bibliometric methods investigate the blurred frontier, such as the study of patents from researchers (Meyer et al., 2003) and conversely the scientific publications from industrial firms (Hicks, 1995). This cross-activity generally results in observable co-activity publication/patents of individuals or institutions, briefly mentioned below.

At the micro level fine grain measures able to capture the relation between a publication and a patent help to depict the scientific neighbourhood of an invention, and the possible range of the technological relevance of a publication. This has an obvious interest for S&T watch and also, at earlier stages, for priority search. Commercial operators tend to offer more and more micro level navigable links between patent and publication databases.

[1] Results are partly based on the experiment presented at the 6th S&T indicators conference (Bassecoulard E et al., 2000)

At the meso level a measurable relation allows one to investigate knowledge transfers and potential spillovers: describing the knowledge base of particular technologies, conversely to disclose the technological neighbourhood of scientific themes/research fronts, the migration of topics in the innovation process, for example from science to technology and society, etc.

At macro level it provides a tool for building concordance tables between scientific classifications and patent nomenclatures and investigating dependences, a helpful instrument for S&T policy.

2.2 A Few Ways of Measurement

The linkage may be investigated by a variety of tools based on the informetric structure of publications and patents: citations, co-activity, classification relations, shared topics (lexical approaches). The first two are analysed in the chapter by Tijssen (op. cit.).

a) *Citation based methods*. They exploit the 'references' field of patents or publications. The analysis of references to non-patent literature in patents is now considered as the classic way of investigation, since the pioneering works of CHI-Research (Narin, 1985) establishing basic patent indicators such as scientific intensity and immediacy of connection to the science base. The usage of patent citations in the economics of S&T and economic geography is now a current practice, since the path breaking work of Jaffe (1988) on local spillovers academe–industry, followed by many others. Extensive works conducted by Schmoch (1997), and recently by Verbeek et al. (2002), lead at the macro scale, to operational correspondence tables between scientific and technological nomenclatures. The symmetrical marker, citations from science to patents, has been systematically surveyed by Glaenzel and Meyer (2003). Patenting publication delays may cause some silence in citations to patents in publications, and conversely if restriction to publish results in extra delays. A good indicator of interaction might combine the two directions of citations. At the micro level the patent–publication citation is relatively easy to implement as a navigational link (for example, ISI-WPI Derwent).
b) *Co-activity methods*. In areas with a particular intensity of S-T exchanges researchers are likely to be involved in both activities. The co-activity may take many forms. Scientists may take patents through their university, through industrial partners, or personally. They may themselves have only academic affiliations or be part time academe/industry. Some of these configurations have been investigated

by Meyer et al. (op. cit). The study of co-activity of scientists has been pioneered by Rabeharisoa (1992). Co-activity studies of academic institutions and industrial organisations are increasingly used (for example, Tijssen and Korevaar, 1997) as a part of the booming area of research on university–industry connections, especially knowledge transfer from universities in multiple forms, contractual or spillovers.

c) *Category sharing in common classification schemes*. Classification schemes of science and technology are basically not commensurable. Whereas patent classifications belong to the strongly codified framework of patent offices, with, in fact, two basic official nomenclatures, USPTO and International Patent Classification (IPC) and variants, there is no commonly accepted scientific nomenclature, so that in bibliometric practice databases' classifications are mostly used: ISI's journal lists and INIST Pascal classification scheme for multidisciplinary databases; classifications schemes for specialised databases. Generally speaking, classifications of publications and patents are not connected with each other. However, there are some (very limited) ready made solutions offered by thematic databases which encompass the two types of objects and propose the same classification scheme for both. The best example of a scientific database with a large coverage of patents is CA of *Chemical Abstracts* Service. However, the classification scheme is coarse grained, so that any in depth correspondence should be made by some other means. Any common classification scheme allowing multiple postings for an item can also be exploited by co-classification studies.

d) *Topic sharing: the lexical way*. At the micro level the topic sharing principle is quite simple. The standard lexical query system implemented in most databases and web engines can be used to retrieve jointly, on a particular subject, publications and patents if present in the source(s). A publication and a patent are retrieved together if their respective sharing of terms with the query is large with respect to the particular settings of the query engine (IR-Information Retrieval formulae/ weighting). In a more general perspective, either in bibliometrics or in cluster models of IR, direct topic sharing between documents can be studied by various techniques. An analogy in the citation world is Kessler's bibliographic coupling (Kessler, 1963). Numerous methods or toolkits for data and text mining of S&T information have been developed with promising markets in mind. Patent offices also aim at providing efficient tools for patent information both for the final user on Internet and for their examiners (e.g. at EPO, ePatent project, 2002; bSmart presented in Sarasua, 2000). Some applications are focused on the patent–publication linkage, for example the non-patent prior art research, others are meant for global S&T watch (Kostoff, 2003), taxonomy development or knowledge discovery, e.g., following Swanson's (1986) approach (Weeber et al., 2000). At the meso level co-word maps mixing publication and patent data have been used by Engelsman and van Raan (1994) to characterise the 'S&T interface' on the basis of visual analysis of multidimensional maps. The question is whether this topic sharing approach can be operationalised in a systematic way and in a large range of scales by calculating a direct coupling between publications (or sets of publications) and patents (or sets of patents). Section 3 addresses related issues.

e) *Hybrid approaches*. Most bibliometricians adopt a pragmatic view and are prone to combine several informetric ways. At the macro level verbalisation of IPC has been investigated by Turner et al. (1991). Conversely, pre-classification by linguistic technologies has been experimented with at EPO (Krier and Zacca, 2002). Faucompré et al. (1997) extended this logic to the publication–patent linkage by using EPO catchwords to re-index automatically the scientific publications in the INIST–PASCAL database. From a different perspective Leydesdorff (2002) linked titles of patents and titles of the scientific documents they cite. Murray (2002) started from patent–paper pairs transcribing the same idea to develop a systematic analysis of the inter-relationship between scientific and technological networks, combining bibliometrics and qualitative methods.

In Section 5 we will discuss some further differences between these competing and complementary approaches.

3. LEXICAL DESCRIPTION OF PUBLICATIONS AND PATENTS AND RELATED SIMILARITY MEASURES

3.1 Informative Features in Different Contexts

In both science and technology, examination systems assess the newness of results claimed by authors/inventors and priorities issues are met in science (Merton, 1957) as well as in technology. Naturally, intellectual property regimes differ deeply: opposition and sanction systems and legal basis are highly codified in technology, whereas informal procedures and a weak legal basis (except copyright matters) prevail on the science side, as, for example, Granstrand (1999) pointed out. One can expect, therefore, different patterns of information disclosure in scientific publications and patents.

Methods and results published in academic journals must be informative enough firstly to convince referees, secondly to attract a large scope of potential users and citers as soon as possible. This is also true for self-standing article surrogates found in bibliographic databases: abstracts, for example, can be considered as 'visiting cards'.

Patent applicants look for financial returns on their research legally warranted by temporary protection in exchange for compulsory disclosure of technical features of the invention. Patent is both a legal document and a piece of technical literature. But a too precise description can be too informative for competitors and is also likely to narrow the scope of the invention (e.g., Gordon and Cookfair, 2000; Sarasua and Corremans, 2000). This shapes well known peculiarities of patent documents; for example, poorly informative original titles.

With the emergence of a knowledge based economy, more and more research teams, academic or industrial, produce results which can be both published in academic journals and patented. Disclosure issues vary between patent systems: the grace delay effective in the US system is still under discussion in the European system. Defensive publication can be chosen to prevent patenting (see, for example, Research Disclosure Journal archives, RDISCLOSURE@ database). In other cases a trade off must be found in schedules and contents of scientific communication media (conferences, publications) and patents, to ensure novelty at the application date and prevent clashes in industry/university collaborations. As quoted in technical on-line brochures of the *European Patent Office* (1998) "publication in a scientific periodical can sometimes give more prior art information than a patent application, particularly where the inventor is writing about his own invention".

As a result of different objectives the respective lexical content of publication and patents, at least in the original documents, may exhibit profound differences. Further processing by databases may reduce this discrepancy.

3.2 Sources

For patents, as well as publications, the usual requirements of data selection for bibliometric purposes apply: comprehensive and relevant data have to be retrieved and downloaded under time and cost constraints. These data have to be machine readable to allow further content processing, and meet basic prerequisites for statistical treatments. This means, to be realistic, that one has to rely on bibliographic databases.

The initial step of the patent–publication relation, the selection of relevant documents, depends on the research question or the operational demand. There is much difference between a meso scale study focused on ST convergence in a geographic area, a micro study of the science base of a given technology, and a macro scale analysis aiming at a correspondence table between a scientific nomenclature and a technology nomenclature. In the first case a geographical codification of actors is sufficient. In the second case a careful delineation of the perimeter is needed (see the chapter by Hinze and Schmoch, Section V in this handbook). In the third case queries have to be designed to optimise processing of large bulks of data. In that case direct agreements with database producers are required, both for legal and practical reasons.

For large scale topic sharing studies by statistical means prerequisites should be considered: redundancy has to be reduced; contents representations of publications and patents should be as homogeneous as possible.

In scientific sources redundancy is a recurrent question which has to be dealt with in some databases or when combining sources. In patent sources the use of basic patents or 'patent families' is advisable, through the unique index in large patent system (US, EPO, PCT), or in specialised databases (Derwent, CA–*Chemical Abstracts*); in other cases it is necessary to process the file for reducing to patent families.

Patent systems and commercial databases of publications and/or patents have various policies in terms of added value in content (indexing/coding; reprocessing title and abstract of patents). In addition, efficient tools for information search and information analysis as well as easily readable export formats must be carefully considered. Last but not least, commercial databases can be very expensive to access.

The added value of the database partly determines the type and amount of pre-processing that will be necessary. Pre-processing (selection, downloading, extraction, and representation of contents, data storage organization) will be generally reduced if data come from traditional databases and formats. Addressing full texts is another challenge.

3.3 Extraction and Representation of Contents

Topic sharing of publications and patents can be assessed by various proximity measures between individual documents and/or sets of documents of the two kinds. At first document contents have to be extracted and represented in ways amenable to quantitative processing. Various combinations of linguistic and statistical methods are used originating in Computational Linguistics, Information Retrieval, and Natural Language Processing, with possible labour division between database, dynamic interfaces, and further analysis by bibliometricians/users. Processing generally involves linguistic unification of terms, and possibly more drastic reduction of dimensions, resulting in a final indexing by broad terms or concepts. Reducing dimensions is required both for convenient handling and reducing silences that result from uncertainty of description and synonymy at large in natural language. Techniques used in data and text mining are presented in the chapter by Leopold et al., Section II in this handbook. We will focus on features related to informetric properties of S&T documents.

Key terms and abstracts, the most usual representations of textual objects, are traditionally added to the source text either by authors/inventors or by the professional indexers/examiners. In bibliographic databases these representations stand as surrogates for the original full texts along with document titles. However, purely human indexing tends to be assisted or substituted by computerised methods. To face information retrieval tasks and reduce manpower costs research into automatic indexing and abstracting was launched in the sixties (Luhn, 1957, 1958). An extensive state of the art report on automatic indexing and abstracting can be found in Moens (2000).

Manageable global representations of publication and patent contents (abstracts in a broad sense) would be likely to open new ways of study of S&T linkages, but their automatic and reliable generation is still largely a research issue. In operational contexts of bibliometrics pre-processing of texts to obtain content descriptors (indexing in a broad sense) seems more affordable.

Free indexing selects significant natural language terms from the texts. The process usually involves recognition and minimal unification of words, proper name recognition, removal of stop words, stemming, and possibly phrases/multi-words recognition and normalisation. Resulting index terms can be weighted according to their importance in the texts and in the whole document collection. Statistical techniques and linguistic knowledge can be involved in term extraction and detection of synonymies or quasi-synonymies. Potential ambiguities (homonymy, polysemy) are handled with more difficulty especially in broad subjects, and word sense disambiguation may call for sophisticated data analysis or graph techniques. Extreme specificities may hinder further relevant document grouping.

Controlled indexing establishes conceptual links between documents and items chosen in an authority list derived from a knowledge base (thesaurus–ontology) of controlled language index terms. A controlled term is the unique form assigned to terms that have similar or related meanings but unrelated surface forms (equivalence class of the controlled term, including synonyms). A thesaurus usually allows one to broaden or narrow the terms that represent concepts found in texts (hierarchical relationship of generic to specific index terms). But it is rather inflexible (regular updates of knowledge bases are needed to account for changes in interests and concepts) and much less portable across different document domains. Automatic building of ontologies is an active research field.

The comparison of automatic and human indexing has been studied for decades (Salton, 1969), and has been renewed by the introduction of sophisticated natural language processing techniques in information retrieval. In an experiment on machine aided indexing Jacquemin et al. (2002) found that a free indexing process, noisier but never silent, proved to be a valuable way of updating a thesaurus when no controlled term (or variant) were available to index a document.

A major problem in S&T literature is Automatic Term Recognition, especially Acronym Recognition, be it polymers, diseases, proteins... Kostoff (2003) gives a particularly spectacular example in a test query of 'IR' in the SCI source database. The problem occurs in all disciplines despite efforts of standardisation in the respective communities, and is particularly severe in chemistry (Chowdhury and Lynch, 1992). The lack of clear naming standards raises the classic issues of term ambiguity (polysemic terms) with negative effects on precision, and term variation (different terms that refer to the same concept) which threaten recall (see, for example, Nenadic et al., 2002).

Task selection, order, and tuning in automatic indexing are still a matter of debate, as is the comparative efficiency of statistical and linguistic techniques (De Bruijn and Martin, 2002). Though according to Sparck Jones (1999) only high level IR tasks — amongst which information extraction or automatic abstracting — could benefit from natural language processing, most practitioners seem to integrate NLP and more traditional statistical approaches in composite text processing sequences. The risk attached to complex chaining of actions may be a black box effect.

To summarise, the reduction of original texts — publications or patents — to manageable representations of contents can be achieved by statistical, morpho-syntactic and semantic treatments combined in human, automatic, or machine aided processes. The reduction can be 'frozen' in permanent fields of the database or dynamically created by dedicated or general engines. These processes are not neutral vis-à-vis informetric properties of content markers (descriptors in a broad sense) used in the following statistical treatments. It is up to the analyst, depending on the source/objectives, to accept the extraction or to re-process reduced forms from original texts. In the experiment reported in Section 4 we took advantage of the good quality of indexing in the CA database to limit further elaboration.

3.4 Statistical Toolbox to Reduce Dimensionality

This objective can be achieved by statistical methods or other means, such as, for example semantic analysis. We focus here on the statistical toolbox.

In the standard case working datasets can be represented in a vector space model (Salton, 1968) where, in the most simple form adapted to keyword descriptions a cell (i,j) has a value set to 1 when Lexical Unit i appears in document j and to 0 otherwise. For natural language texts the cell value is typically set to the intra-text frequency of the term. Depending on the technical requirements and the methods, the starting point is either two rectangular matrices [N1,M1] for articles and [N2,M2] for patents, or a single one [N,M], where N = (N1 + N2), the total number of documents, and M is a common repertoire of Lexical Units, the union of M1 and M2. The second option is more in line with further reduction of dimensions towards a common scheme.

The similarity between a patent text and a publication text is a particular case of inter-text similarity. As mentioned before, the specificity of technical language and patent jargon on the one hand, and publication rhetoric on the other hand, may create false distinctions between Lexical Units (probably more than false equivalences) if 'translation' problems, in the broad meaning conveyed by the 'sociology of translation', are not addressed. False distinctions will result in silences when calculating patent–publication relations. A controlled vocabulary or the reduction to unified terms or concepts is expected to bring partial solutions.

3.4.1 Similarity

Similarity between lexical units or texts has received attention from several disciplines. Firstly, a specialised area of statistics and data analysis deals with texts' properties, terms' distribution, allowing for a similarity calculation. Correspondence analysis, for example, was early targeted at lexical studies (Benzecri, 1981). Secondly, a particular case of similarity between two lexical forms, a query and a document description, is central in Information Retrieval. In the standard IR 'vector space model', similarity between query and documents is used to rank documents by relevance. Similarity between documents is assessed in a 'lexical coupling' rationale kindred to query–document similarity, the proximity of two documents depends primarily on the number of Lexical Units they share.

Amongst the vast choice of classic Euclidian and non-Euclidian measures, some have been privileged in IR and textual statistics. Common statistical measures are often adapted either to weight lexical units as a function of frequency, or to deal with the effects of texts' length. Most usual similarity indices are cosine measures, Jaccard, Dice. Such local measures have been advocated, for example, by the promoters of co-word in sociology of innovation (Callon et al., 1986b). Probability indices (observed linkage over expected linkage) highlight weak signals. Euclidian proximity is also used.

Some weighted forms are common, for example the 'best fully weighted system' where the cosine function is weighted by the TDF-IDF score, proposed by Salton and Mc Gill (1983) (see also the chapter by Leopold et al., op. cit.). Amongst Euclidian distances a powerful form of weighting is the Chi-Squared on which correspondence analysis is based. Chi-Squared weighting favours low–frequency items, and exhibits the property of distributional equivalence (neutrality of aggregation of items with similar profiles). These measures can be applied to Lexical Units or documents. Specific metrics couple measures in the two universes (factor analysis, below).

Further comparison of distributions of terms between two texts can be transposed from the query document comparisons investigated by probabilistic models of IR (see below), for example in Bookstein and Swanson (1974).

3.4.2 Dimensionality reduction

The objective is now to obtain a matrix [N, L] where L is the number of final structuring items, for example concepts, document clusters — or directly a [N1, N2] matrix of similarity after rows or columns reduction. The transformation is obtained by a reduction of dimensions involving similarity calculation and grouping of documents or, as a possible intermediary, of Lexical Units. Reduction is generally more drastic than unification of linguistic forms and keyword indexing mentioned above.

As often in bibliometric applications, for example themes mapping, several ways are offered: starting with rows (grouping documents); with columns (grouping structuring items); or using dual methods dealing simultaneously with the two universes.

Most dimensional reduction techniques fall into either the factor or the clustering family (see also the chapter by Leopold et al., op. cit.).

– *Factor family*. Factor analyses are continuous techniques that disclose latent variables as combinations of original variables. They work simultaneously in documents and items universes. Factor loading can be used to re-index documents by latent variables; for example, to provide a robust re-indexing common to patents and articles. Correspondence analysis (Benzecri, op. cit.) is particularly interesting for textual analysis, with a symmetrical treatment of individuals and variables. With a different metrics, the 'Latent Semantic Analysis' (LSI, Deerwester et al., 1990) is increasingly used in IR and data mining contexts.

– *Clustering family*. Clustering methods are discontinuous methods quite common in bibliometrics and IR, for Lexical Units as well as for texts. Co-citation research fronts and co-word themes on the one hand, bibliographic coupling clusters based on citation (Kessler, op. cit.) or on words on the other hand are classic examples. IR makes use of cluster models, with pre-processed clusters or interactively built clusters. Recent neuronal and other Artificial Intelligence approaches are presented in Kiang (2003). Multilevel algorithms, ascending or descending, suppose a definition of (dis)aggregation criteria and hence inter-cluster distance.

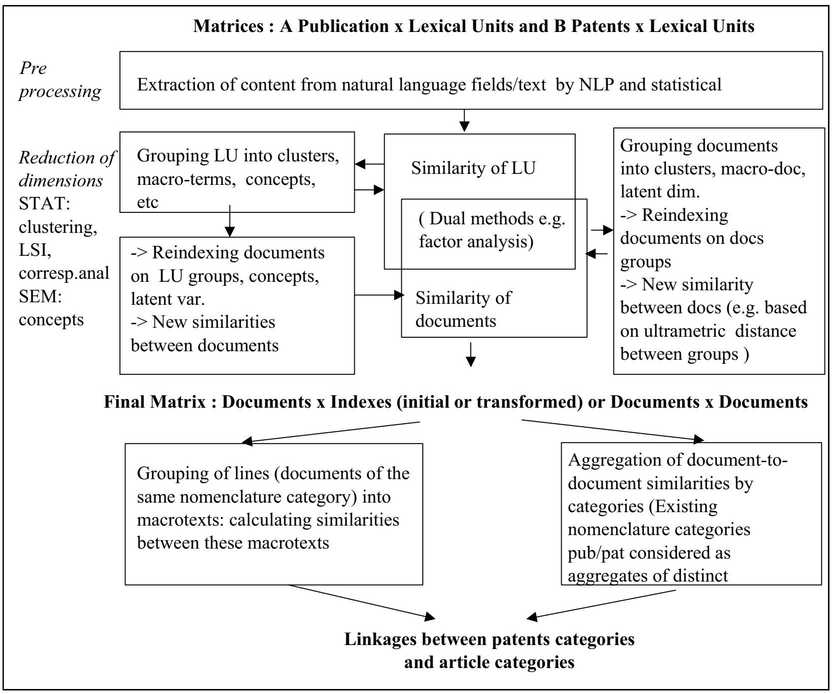


Figure 30.1. From texts to S&T aggregates linkages

Factor analysis is generally claimed to be more powerful since it reveals latent dimensions and directly allows overlaps (as do some clustering techniques). On the other hand when high dimensionality levels are required for output (several hundreds), clustering techniques can be more effective. In fact combination of clustering and factor analysis in either sequence is common in data analysis. Some couples of techniques are consistent (Ward clustering/ correspondence analysis). Current development of data mining fosters availability and tractability of methods (neuronal clustering, fast-clustering, refined sampling, cross-checking of methods, etc.).

A typical scheme is summarised in Figure 30.1. (LU stands for Lexical Units).

3.4.3 Final matrix — application to similarity between *ex ante* aggregates

Depending on the method, the reduction leads either to a new coordinate space with reduced number of columns or to a square table of similarities. For example, in the case in which documents have been clustered by a hierarchical method, the similarity matrix can reflect the cut off of the tree at a given level, with either a Boolean recoding of similarity between items (1 if the two documents belong to the same cluster, else 0), or a new similarity s, with $0 \leq s < 1$ (1 if the two documents belong to the same cluster, else a value derived from the ultrametric distance between their clusters). In the case in which terms have been clustered in macro terms a new similarity between documents may be calculated using this re-indexing, etc.

At the end of the reduction process, or without reduction if a satisfactory controlled language is already available, the proximity between a patent and an article (and if required, proximity between patents or between articles) can be assessed on a more robust way than before reduction.

This can be used for a variety of applications, as mentioned before. Let us only consider the correspondence between nomenclatures, which is a particular case of inter-group similarity assessment. Nomenclature categories are defined *ex ante*, say IPC categories of some level, and individual journals or set of journals. Two strategies may be used to assess inter-category proximity:

a) First group documents, then calculate similarity between groups' articles, respectively patents, belonging to the same nomenclature category are merged into macro texts with aggregate coordinates, typically on the [N, L] matrix. Then similarities between these macro texts are evaluated.
b) First calculate document to document similarity, then aggregate similarity.

The starting point is the final matrix of similarity between individual articles and individual patents. The individual linkages are then aggregated at the category level on particular rules.

The first strategy has low computer requirements, since (at least at this stage) item to item distances are not needed. Aggregation further reduces silence. Advantages of the second strategy: simpler calculations; keeping information at the micro level and the related optimisation means (discarding non-significant linkages); flexibility of aggregation whatever the scale. Citation linkages are Boolean: document A do or do not cite document B. In contrast, lexical linkages between documents used in the second strategy (b) range in a continuum: from complete discordance (nothing in common) to complete concordance (same collection of content markers/descriptors). A Boolean transformation using a threshold may be practiced.

In the experiment reported below, thresholds are applied to item to item similarity, resulting in a Boolean measure of the publication–patent linkage. Whether Boolean or not, the same variants as for inter-cluster distances in classification algorithms are offered. A rationale of average linkage seems prudent, with optional transformations (log) limiting instability owed to discrepancies in category coverage.

Starting from a Boolean measure of the article patent similarity, simple measures of connection between a science category and an IPC category are: the gross number of connections; the ratio of observed connections to all possible connections (average linkage); the product of proportions of connected items on both sides (mutual inclusion or Equivalence Index)

3.4.4 A few limitations

Similarity based models have been criticised for lacking theoretical anchorage and a large literature has been devoted since the seventies to alternative models of IR based on probability. The assumption that representations are uncertain (for example, indexing is not an exact science) leads to key concepts of probability of relevance and probability ranking principle. Probabilistic IR has developed in many directions: model oriented approaches, description oriented approaches (see the review by Crestani et al., 1998). At a high level of generality, vector space and probabilistic models are, however, not antagonist. The probabilistic inference model (Wong and Yao, 1995) unifies most IR models into a general approach based on a concept space in which a query and a document represent a need and a content of knowledge. In this model probabilities are based on semantic relations. Users of factor analyses, although these techniques are statistical rather than semantic, would probably recognise a concept space as an acceptable framework for their quest.

Besides, models based on comparison of lexical lists are not universal. Without mentioning other media challenging IR (images, videos, etc.), text–based documents may be represented by semantic structures rather than lists:

set of logical rules; conceptual graphs; etc. In these cases the distance between document A and B asks for more sophisticated approaches. A general statement is that distance is proportional to the number of steps needed in a given protocol to match the representation of B by successive transformations of A. We will not expand on these models, hardly manageable up to now for large bibliometric applications, but that could gain diffusion in the future.

4. ANALYSIS ON AN EXAMPLE

4.1 Data

In this experimentation we intended to assess the feasibility of a lexical approach to the issue of a science–technology nomenclature concordance, here a correspondence between technical fields (IPC) and scientific specialties (journal–based), namely ISI-SCI subject category codes.

We have seen in section 2 that, ideally, the exploration should be carried on a multidisciplinary scientific database and patent databases including the most important patent systems. In addition, to guarantee a safe statistical basis, contents of patents and articles have to be represented in a homogeneous, reliable way, which requires heavy and sophisticated language processing tasks.

Not to face 'all problems at once' we chose a database collecting both articles and patents in an homogeneous way, *Chemical Abstracts* produced by the American Chemical Society, which covers chemistry related literature (ca 9,000 journals, conference proceedings, dissertations, technical reports, books) and patents from several granting systems, both national and international. Among the ca. 750,000 documents retrieved per year, journal articles account for 74% and patents for ca 16%, so that patent examiners use CA databases in the chemical area for much of their searching (Gordon and Cookfair, op. cit.; EPO (1998) online information). In fact, CA coverage extends to other physical sciences, and also to life sciences through biochemistry and biophysics.

Patent coverage of CA is selective, but exhaustiveness is guaranteed for certain groups within a list of IPC subclasses, available on CA documentation online. Some subclasses are completely covered, especially in chemistry and metallurgy. The trial has been carried out on a sub-sample of CA database, covered in a thematic CD-ROM[2] 'Food & Feed Chemistry'. Inside this coverage, the variety of subjects has been kept: food science and technology, nutrition and cancer, toxicology, fats and oils, cosmetics. Not all groups of relevant IPC subclasses are covered in this particular source.

Detailed technicalities will be published in a forthcoming article. Some results of processing steps are gathered in Tables 30.1 and 30.2 below.

22,969 articles and 3,855 patents from the four main granting systems (USPO, JPO, EPO, WIPO-PCT) have been extracted from the CD for publication year 1997. Patents were assigned only to the main IPC class. Beside bibliographic information or patent information, indications of topics covered are found in three different types of items of CA database available both for patents and articles:

a) natural language: title, abstract, authors keywords;
b) controlled terms: general subject index entries;
c) specific terms: CA Registry numbers, unique identifiers of chemical substances.

Automatic inclusion of generic terms in the controlled vocabulary has been kept. The advantage is to retain possible coupling that would disappear as a result of very detailed indexing. A serious shortcoming is the multiplication of forms that unduly enhance linkages by creating spurious connections from a single shared term. On both controlled and natural vocabularies, hapaxes (terms that occur only once) have been discarded as they cannot generate a lexical coupling, so that only 20,669 articles and 3,853 patents remained in the working dataset.

In the following, 'controlled vocabulary' refers both to controlled index terms and to CA registry numbers, the latter very valuable given the difficulty of Automatic Recognition of chemical substances. This controlled vocabulary is not totally bias-free, because CA focus on chemistry related aspects: medical or mechanical notions, for example, may be undervalued. Nevertheless the overall quality is expected to be good and to correctly emulate a reliable content extraction by sophisticated language processing techniques, some of them perhaps used as auxiliaries by the database.

[2] Chemical Abstracts lend us the thematic CD ROM Food & Feed Chemistry, VOLUME 1999, Issue 12, to carry on the experimentation (ca. 30,000 documents per publication year, 23,000 articles and 4,700 patents). Some fields available on-line, such as 'roles' and codes of the 80--section classification scheme, used for example by Morillo (2001), were not included in this version of the database.

The experiment on natural language was limited to a basic control of natural language terms of titles and keywords[3] without any attempt to process abstracts and extract concepts by reduction of dimensionality. This is meant to figure a reference of 'worst case', with poor native information and extraction process, in contrast with the expected high quality of controlled vocabulary.

Frequencies of the two types of descriptors in articles and patents are shown on Table 30.1. The dictionary of natural terms is 50% larger than the dictionary of controlled terms (after discarding the hapaxes). The number of shared terms (found in both types of documents) is quite similar, resulting in a higher share for controlled descriptors (46%) than for natural terms (30%).

Table 30.1. Controlled and natural descriptors in the dataset

Type of Vocabulary	*Number of distinct descriptors*	*Mean (Maximum) frequency*	*Distribution*		
			Articles only	*Patents only*	*Common*
Controlled	15,445	17 (3183)	44%	10%	46%
Natural	22,932	15 (4234)	64%	6%	30%

4.2 Measuring Document Similarity

The principle is a document to document measure, the only one totally flexible for various aggregations. Usual similarity measures involve the number and/or the proportion of terms in common. We choose a set of Jaccard indices, in order to allow length normalisation and term weighting.

Medium frequency terms are often considered as carrying the most significant information. "When term independence and binary indexing are initially assumed, the most important terms exhibit medium frequency and the worst ones are the high frequency terms" (Salton and Wu, 1981). To favour terms of intermediate frequencies (although total independence cannot be assumed here because of automatic generation of generic terms), we have used a log based 'parabolic' weighting scheme assigning a maximum weight to words of mean frequency and a minimum weight both to words of minimum frequency and words at an upper frequency threshold. To remove poorly informative terms the weight is set to 0 above the upper frequency threshold.

[3] Extraction of words and multi-words using stop words, standardisation of abbreviations, and unification of most singular-plural forms achieved by Perl programs (D. Besagni, INIST). Manual inspection of resulting terms.

Because very long lists of descriptors can be encountered, for example for chemical substances, we used a simple log transformation to minimise the length problem. For patent x and article y, let n_x and n_y be the number of their index terms, n_c the number of shared descriptors, p_x, p_y, p_c the corresponding sums of weights.

The plain Jaccard index is the ratio of (n_c), intersection of the two lists of descriptors to their union ($n_x + n_y - n_c$). The weighted Jaccard index will be the ratio of the corresponding sums of weights. Both indices can be calculated without or with log length normalisation (here by taking the logarithms of the intersection and the union).

For example, we used following weighted Jaccard index with length normalization:

$$\text{Jacwlog} = \log(p_c) / \log(p_x + p_y - p_c)$$

Given the number of articles and patents (more than 70 x 10^6 possible pairs) we choose to keep as valid links only those involving at least 3 words and 1 common 'informative' word (non-zero weighting). This first selection is all the more necessary because of the automatic addition of generic terms mentioned above. Selected links represent 0.25% of possible linkages on controlled vocabulary. The fuzzier nature of natural vocabulary leads to a stronger selection, with 0.17% of possible links in this case.

Then we reduced these linkages to Boolean values with a threshold on similarity indices (mean + one Standard Deviation on the index Jaccwlog in following examples), which discards many not significant links, and delineates a more precise neighbourhood for each article or patent. The effect of the threshold on the number of neighbours is shown on Table 30.2.

Table 30.2. Effects of statistical threshold* on proximity index: Average number of linkages per article and patent

Document type	*Controlled vocabulary*		*Natural vocabulary*	
	No threshold	Threshold	No threshold	Threshold
Article	15.6	3.2	9.9	3.3
Patent	71.3	9.5	39.7	8.8

*Mean + 1 STD on Jaccwlog

4.3 Measuring Proximities of Articles and Patent Categories

To allow flexible ways for any aggregation, we chose to aggregate pre-calculated item to item (Boolean) similarity rather than building a macro text for each category and then evaluating similarity between macro texts. We used one of the simplest measures of the linkage between a patent and a science category: the product of proportions of connected items on both sides, termed 'mutual inclusion'. It must be kept in mind that the number of patents in IPC category i and the number of articles in science category j are counted within the dataset, that is with respect to IPC and journal coverage of the our extract of CAS (thematic CD).

A necessary improvement in further studies will concern the processing of automatically added generic terms, which proves to avoid silences in lexical coupling but tends to spuriously multiply the number of linkages. The solution of this problem should be sought at the similarity calculation stage. For convenience results are reported only for those links which connect patents to articles in ISI-covered journals (48% of the journals and 59% of the articles).

4.4 Results

A first outcome is in terms of IR features. The number of neighbours for a given patent (Table 30.2) is about 10 for a sensible threshold corresponding to a pretty strong selection. This confirms the expected position of the lexical approach compared to the citation approach. Michel and Bettels (2001) record an average inferior to one non-patent reference per patent in the European system for publication year 1999, and about three times more in the USPTO. The relevance of each neighbour detected by lexical connection would be to check individually to secure interpretation in terms of recall and precision, but a first manual checking confirms the reality of a 'topic sharing' in randomly examined cases. The 'topic' may be of different nature and reflect some kinship in products, substances, processes, in some cases models or methods. As mentioned above, some improvements can be made to select significant linkages, but one cannot expect that the lexical linkage at the micro level can detect article patent couples sharing a same research question. Much higher thresholds could perhaps approach this objective at the expense of recall, but clearly better precision can be expected from citation linkages.

Figure 30.2 compares the outcome of various settings to depict the science base of a wide scope IPC subclass (Miscellaneous Food preparation and preservation). Scientific categories are ordered by the rank of the expected best measure that uses threshold on controlled vocabulary. Removing the threshold on controlled vocabulary does not alter the ranking of the three first science categories (Food Science, Applied Chemistry, Agriculture), but bring some rank shifts afterwards. The comparison between a 'quick and dirty' processing of natural language on the one hand and controlled vocabulary on the other hand is quite interesting. Especially when using the threshold, the discrepancy between rankings on natural and controlled language is not as high as could have been expected. This is an argument in favour of a certain robustness of the lexical approach.

With the expected best option, controlled vocabulary with threshold, IPC A23L is connected first to Food Science, followed by Applied Chemistry, Agriculture, and Analytical Chemistry. The scientific dependence of specialised IPC subclasses is not as scattered. For example A23C Dairy (not shown) is firstly connected to Food Science, and then four categories: Applied Chemistry, Analytical Chemistry, Agriculture-Dairy and Horticulture.

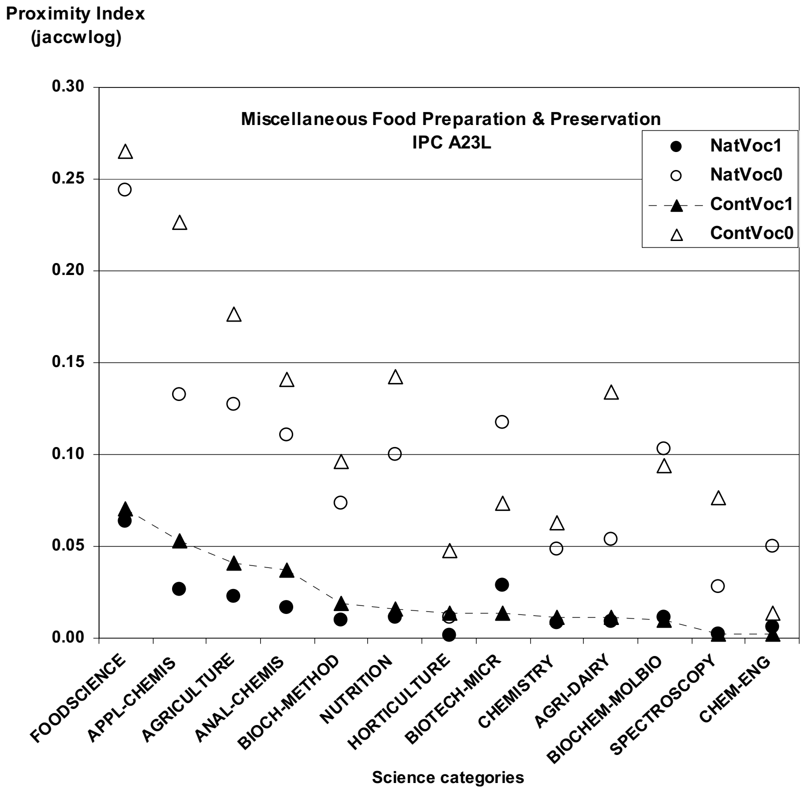


Figure 30.2. Pattern of the science base of an IPC subclass: effect of methodological options

NatVoc1: natural language, threshold ContVoc1: controlled language, threshold
NatVoc0: natural language, no threshold ContVoc0: controlled language, no threshold

Results should be carefully interpreted, if only because of the coverage arbitrarily limited to the contents of the source (thematic CD) and of perfectible methodology, especially the treatment of automatic generated terms. Figure 30.3 shows a tentative correspondence table in the area examined.

Categories are ranked, on both sides, by their median linkage with other side categories. Low linkage categories are not represented (some may exhibit isolated strong connection with a few particular partner category). The CD coverage may cause important linkages to be obscured, and it should be remembered that all groups are not covered in the sub-classes under examination.

Concentration of linkages amongst privileged categories and pairs of categories is clear, but perhaps somewhat reinforced by the particular measure of proximity chosen.

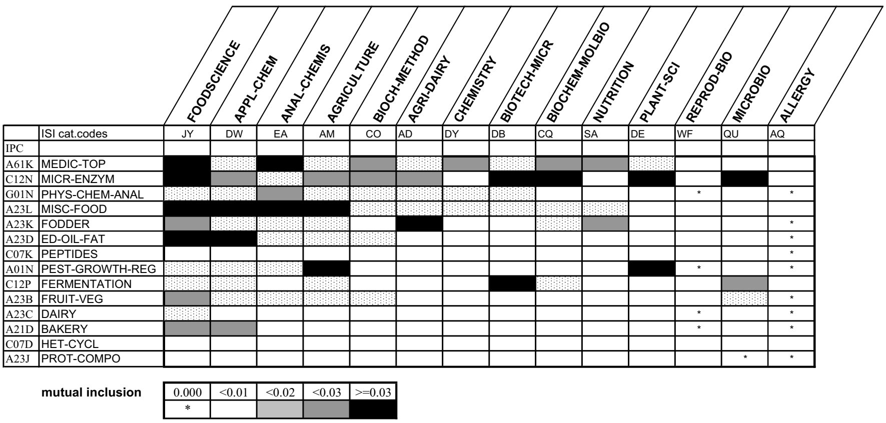


Figure 30.3. Tentative correspondence table (extract) based on mutual inclusion index

In spite of these reservations, two main conclusions can be put forth:

a) Lexical connection is definitely exploitable at the macro level to establish concordance tables.
b) Over-interpretation should be avoided at the micro level (one to one coupling).

5. DISCUSSION AND CONCLUSION

5.1 Lexical *versus* Other Bibliometric Techniques

5.1.1 Nature of the linkage

Let us summarise some technical properties of the methods mentioned above. For memory's sake category sharing is a very effective indication based on indexers' skill, but information is rarely available, and is constrained in precision by the level of breakdown nomenclatures (for example in CA).

Co-activity is relatively simple to define and capture, at least for personal co-activity, with respect to the usual precautions in bibliometrics for unification of personal and institutional names. The publication–patent relation created by co-activity relation is symmetrical, and dynamic interpretation (for example, antecedence science–technology) is possible with respect to careful examination of technical delays (patent publication cycle, grace period if applicable, publication delay, etc.).

Citation linkage and lexical/linguistic linkage raise perhaps more difficult issues. Their difference is well known in bibliometrics. The citation linkage is explicit, voluntary, selective and asymmetrical. The time difference between citing and cited article is interpretable (at least in a Mertonian scheme) as a delay in the use of a former knowledge in a dependence context. The act of citing is strongly selective: only a few articles amongst citable ones are chosen, and often in a reduced repertoire, as many studies of skew distributions and Matthew effects have shown (e.g. Price, 1976). As demonstrated by Narin (1994, op. cit.), formal properties of citations in patents are quite similar, although their nature and interpretation are deeply different.

In contrast, the lexical relation is implicit, symmetrical, achronic. It is not established directly but disclosed by the comparison process of two texts or lists of markers. The interpretation of the linkage is in terms of topic sharing or social relation, rather than dependence (a quite large sociological literature has been devoted to the role of texts in science, see Callon et al., 1986a). The achronicity of the lexical relation does not allow one to detect dependence at the micro level. At the macro level dynamic interpretations of the drift or expansion of vocabulary in given fields, analysis of transfer of vocabulary from publications to patents, or from both to professional media and society are promising, but beyond the scope of this chapter.

Although based also on selection processes within technical and social repertoires, the lexical relation appears less scarce than citation. A topic sharing analysis should be able most of the time to recover larger sets of 'citable' objects and not only cited ones. Whether the lexical approach can operationalise this potential of topic sharing analysis depends on many factors and settings. Statistical distributions of items and linkages help to summarise the expected conditions of recall and precision.

5.1.2 Expected IR properties

Common science–technology nomenclatures, seldom available (CA is amongst the exceptions) are coarse grained only. Co-activity at the personal level is a very high precision and very low recall technique, only able to measure linkages within areas showing outstanding S&T interaction (for example NTIC and biotechnology).

The comparison between citations and lexical techniques is enlightened by the typical statistical properties of the two universes, which can be studied:

– By usual aggregate distributions. Lexical distributions are generally described as hyperbolic Estoup–Zipf–Mandelbrot distributions (Zipf, 1949; Mandelbrot, 1953), extremely concentrated. Citations distributions are also skewed, but at a lesser degree. In this respect the 'citation vocabulary' is richer, i.e., more complex than the language, especially the natural language (controlled language is usually more complex according to this definition). Properties of citation distribution in patents have received less attention than publication distribution properties, until Narin's works (1994, op. cit.).
– By distributions of linkages and graph/social networks properties.
– By direct modelling of recall. A tool for comparing maximum expected recalls in an ideal type of Boolean retrieval scheme is the 'referencing structure' function, which unifies in a disaggregated form citation and reference distributions and can be generalised to lexical distributions for comparison (Zitt et al., 2003).

The distributional properties give some evidence that the IR trade off is different for citation analysis and lexical analysis: high precision and low recall for citations, high recall and low precision for lexical analysis. Lexicology can recall many more linkages, but at the expense of 'false connections', built, for example, by high frequency words and polysemic words. In fact the signal noise properties or the recall precision properties of lexical measure depend firstly on the material (database dependent controlled terms, natural language) and secondly on the statistical settings (type of formula, weighting, thresholds, see Section 3). Their particular properties (recalled in Table 30.3) make this collection of methods clearly complementary.

Table 30.3. Expected properties of various measures

Type of linkage	*Nature of the relation*	*Characteristics of linkage*	*IR properties*	
			*Silence**	*Noise***
Subject sharing (lexical proximity) 1. natural language.	similarity of contents, evaluated: on natural language fields of notices	*source*: pairs of items *relation*: symmetrical, 'achronic' *measure*: real or Boolean (depending on methods)	Low to Med. (varies with setting)	High to Med. (varies with setting)
2. controlled language.	on controlled language fields (high quality databases)	id°	Med. (higher than above)	Med. (lower than above)
Category sharing	similarity of contents, evaluated by the database classification in the same categories	*source* : nomenclature categories *relation*: symmetrical, 'achronic' *measure*: Boolean (basically)	Low (varies with the grain of the classification)	High (varies with the grain of the classify-cation)
Activity sharing (Co-activity 'scientists inventors')		*source*: level of individual players *relation*: symmetrical *measure*: Boolean	Very High	Very Low
Citations from patents to publications Citations from publications to patents	referencing by authors (e.g. scientific background) and indexors (priority context)	*source*: individual patents *relation*: asymmetrical, diachronic *measure*: Boolean	High	Low

* silence = 'true' linkages not found (the lower the silence, the higher the recall)
** noise = linkages unduly detected or 'false' linkages (the lower the noise in the results, the higher the precision)

The high recall of lexical methods is particularly valuable in low signal areas where citation or co-activity measures are unable to capture operational linkages. Their low precision is expected to penalise interpretations at very low level (the item to item relation) rather than at the meso or macro level (for example, nomenclatures' correspondence).

It must be stressed that lexical or category sharing proximity can be studied between any items whatever the dates of publication. This achronic character is an advantage in the concordance studies context. The only citation based method with a similar immediacy, 'bibliographic coupling' (reference sharing) able to establish a linkage between two simultaneous items, is common in science bibliometrics but hardly usable for patents because of too short references lists. In contrast, the classic citation studies examine diachronic relations between today technology and yesterday science.

5.2 Results and Perspectives

The only purpose of the experiment reported was to test the feasibility of a lexical approach to the patent–publication relationship in a favourable case in which a database (CA) records both types of documents in a similar framework. Although CA coverage of scientific literature and patents is quite large, it remains focused on physical sciences and biochemistry/ biophysics and matches neither the complete scope of SCI or Pascal, nor the coverage of patent offices' databases or general ones.

Within the perimeter of CA, patent–publication relation analysis proved feasible using several lexical methods, either based on controlled language fields or on natural language fields. The quality of the controlled language, which includes registry numbers, is remarkable. However, the analysis may be biased as a result of indexing choices made by the database providers, for example. In CA the focus is on chemical aspects and analytical methods. A serious technical problem stems from the automatic posting of generic terms, which can be solved by an adaptation of the similarity calculation algorithm.

Contrary to patent citations to scientific papers, which are readily available and only need standardisation tasks, the lexical relation should be calculated. Moreover, the dimensionality reduction, desirable if natural language is used, may involve fairly heavy processing. If only macro level relations are sought, the process may be alleviated by the comparison of macro texts combining documents corresponding to a publication category and a patent category, respectively. However, document to document proximity data are more informative and allow flexible aggregation.

A large scope of methodological choices is open: choices of indices, of weighting, of thresholds, resulting in various IR properties. Results are highly sensitive to these options, and a wide range of recall/precision trade off can be achieved. As various elaboration stages may take place in the database and in the interfaces, including not documented automatic methods, the analyst should be aware of the risk of artefact.

At the micro level topic sharing is rather well captured by the lexical connections, but rarely at the point at which topic sharing indicates a real dependence in terms of research questions or application path. Citation relation performs better in this respect. Two-stage perspectives are offered:

– Firstly, the systematic study of patent–publication relations in databases, such as CA, which process both items with similar methods. A major limitation is of course the disciplinary coverage.
– Secondly, the extension to non-integrated databases, preferably multidisciplinary (ISI, more stable and INIST–Pascal, with classification scheme and indexer's descriptors) on the science side. On the patent side, the choice is between patent offices' databases and Derwent, the latter with an elaborated classification scheme and lexical fields.

Up to now the multidisciplinary data sources for patents and publications were separated. The trend today is towards an integration of knowledge information whatever the nature. A new sign perhaps is the connection between ISI publications and Derwent patents (two databases belonging to the same company) by navigation along the citation linkages. The navigation between documents along a lexical connection is also likely to expand. However, facilities to aggregate linkages have not been implemented so far and need particular studies. Given the stakes of a better understanding of science–technology relations at all levels, benchmarking analyses of the various forms of informetric linkages (citations, topic sharing, category sharing, co-activity) will be of particular interest in the coming years.

ACKNOWLEDGEMENTS

The authors wish to thank D. Besagni, X. Polanco and B. Maudinas from the Institut de l'Information Scientifique et Technique (INIST–CNRS) for their help in the first experiments on this project, and *Chemical Abstracts* Service for lending us a thematic extract of their database.

REFERENCES

Bassecoulard, E., Polanco, X., Zitt, M. (2000). *Science–technology relationship: the Lexical Connection*. 6th S&T Indicators Conference, 24–27. Leiden: CWTS.

Benzecri, J.P. et coll. (1981). *Pratique de l'analyse des données : Linguistique et lexicologie.* Paris: Dunod.

Bookstein A., Swanson, D.R. (1974). Probabilistic models for automatic indexing. *Journal of the American Society for Information Science*, 25 (5), 312–318.

Callon, M. (1994). Is science a public good? *Science, Technology and Human Values*, 19, 395–424.

Callon, M., Law, J., Rip, A. (1986a). *How to study the force of science*. In M. Callon, J. Law, A. Rip (Eds.), Mapping the dynamics of science and technology (pp. 3–15). London: Macmillan Press.

Callon, M., Law, J., Rip, A. (1986b). *Qualitative scientometrics.* In M. Callon, J. Law, A. Rip (Eds.), Mapping the dynamics of science and technology (pp.107–123). London: Macmillan Press.

Chowdhury, G.G., Lynch, M.F. (1992). Automatic interpretation of the texts of chemical patent abstracts .1. Lexical analysis and categorization. *Journal of Chemical Information and Computer Sciences*, 32, 463–467.

Crestani, F., Lalmas, M., van Rijsbergen, C.J., Campbell I. (1998). Is this document relevant? ... probably: A survey of probabilistic models in information retrieval. *ACM Computing Surveys*, 30 (4), 1–30.

Dasgupta P., David P.A. (1994). Toward a new economics of science. *Research Policy*, 23, 487–521.

David P.A., Foray D. (1995). Accessing and expanding the science and technology knowledge base, *STI Review,* 16, 13–68.

De Bruijn B., Martin, J. (2002). Getting to the ©ore of knowledge: mining biomedical literature. *International Journal of Medical Informatics*, 67, 7–18.

Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41 (6), 391–407.

Engelsman, E.C., van Raan, A.F.J. (1994). A patent based cartography of technology. *Research Policy*, 23, 1–26.

ePatent website : http://www.eu-projects.com/epatent/ (last visited 28/11/2003).

Etzkowitz, H., Leydesdorff, L. (1997). Universities and the global knowledge economy: a triple helix of university–industry–government relations. London: Pinter.

European Patent Office (1998). Organisation of search and documentation in DG 1. http://www.european-patent-office-org/dg1/brochures/index-search-doc.htm. (last visited 28/11/2003).

Faucompré, P., Quoniam, L., Dou, H. (1997). An effective link between science and technology. *Scientometrics*, 40 (3), 465–480.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M. (1994). *The new production of knowledge: the dynamics of science and research in contemporary societies*. London: Sage.

Glaenzel, W., Meyer, M. (2003). Patents cited in the scientific literature: an exploratory study of reverse citation relations. *Scientometrics*, 58 (2), 415–428.

Gordon, T.T., Cookfair, A.S. (2000). *Patent fundamentals for scientists and engineers*. 2nd Edition, Boca Raton (FA): CRC Press.

Granstrand, O. (1999). *The economics and management of intellectual property*. Cheltenham: Edward Elgar.

Grupp, H. (1998). *Foundations of the economics of innovation*. Cheltenham: Edward Elgar.

Hicks, D.M. (1995). Published papers, tacit competencies and corporate management of the public/private character of knowledge. *Industrial and Corporate Change*, 4 (2), 401–424.

Hinze, S., Schmoch, U. (2004). *Opening the black box. Analytical approaches and their impact on the outcome of statistical patent analysis*. In W. Glaenzel, H. Moed, U. Schmoch (Eds.), Handbook of Quantitative Science and Technology Research. Kluwer Academic Publishers.

Jacquemin, C., Daille, B., Royauté, J., Polanco, X. (2002). In vitro evaluation of a program for machine aided indexing. *Information Processing and Management*, 38, 765–792.

Jaffe A. (1989). Real effects of academic research. *American Economic Review*, 79, 957–970.

Kessler, M.M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14, 10–25.

Kiang M. (2003). A comparative assessment of classification methods. *Decision Support Systems*, 441–454.

Krier, M., Zacca, F. (2002). Automatic categorisation applications at the European patent office. *World Patent Information*, 24, 187–196.

Kostoff, R.N. (2003). *Text mining for global technology watch*. In M.A. Drake (Ed.), Encyclopedia of library and information science (pp 2789–2799). New York: M. Dekker.

Leopold, E., May, M., Paass G. (2004). *Data mining and text mining for science and technology research*. In W. Glaenzel, H. Moed, U. Schmoch (Eds.), Handbook of quantitative science and technology research. Kluwer Academic Publishers.

Leydesdorff, L. (2002). Researching the hidden Web: patents and the science base of technologies. http://www.leydesdorff.net/HiddenWeb/HiddenWeb.pdf. (last visited 28/11/ 2003).

Luhn, H.P. (1957). A statistical approach to mechanized encoding and searching of literary information. *IBM Journal of Research and Development*, 1 (4), 309–317.

Luhn, H.P. (1958). The automatic creation of literature abstracts. *IBM Journal of Research and Development*, 2 (2), 159–165.

Mandelbrot, B. (1953). *An information theory of the statistical structure of language*. In W. Jackson (Ed.), Proceedings of the 2nd Symposium on applications of communication theory. London : Butterworths.

Meyer, M., Siniläinen, T., Utecht, J.T. (2003). Towards hybrid Triple Helix indicators. *Scientometrics*, 58 (2), 321–350.

Merton, R.K. (1957). Priorities in scientific discovery: a chapter in the sociology of science. *American Sociological Review*, 22, 635.

Michel, J., Bettels, B. (2001). Patent citation analysis: a closer look at the basic input data from patent search reports. *Scientometrics*, 51 (1), 185–201.

Moens, M.F. (2000). *Automatic indexing and abstracting of document texts*. Kluwer international series on information retrieval. Norwell, MA: Kluwer Academic Publishers.

Morillo, F., Bordons, M., Gómez, I. (2001). An approach to interdisciplinarity through bibliometric indicators. *Scientometrics*, 51 (1), 203–222.

Murray, F. (2002). Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. *Research Policy*, 31 (8–9), 1389–1403.

Narin F., Noma E. (1985). Is technology becoming science? *Scientometrics*, 7, 369–381.

Narin F. (1994). Patent bibliometrics. *Scientometrics*, 30 (1), 147–155.

Nenadic, G., Mima, H., Spasic, I., Ananiadou, S., Tsujii, J. (2002). Terminology driven literature mining and knowledge acquisistion in biomedicine. *International Journal of Medical informatics*, 67, 33–48.

Pavitt, K. (1985). Patent statistics as indicators of innovative activities : possibilities and problems. *Scientometrics*, 7, 77–99.

Price, D.J. de Solla (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27, 292–306.

Rabeharisoa, V. (1992). *A Special Mediation between Science and Technology: When Inventors Publish Scientific Articles in Fuel Cells Research*. In H.Grupp (Ed.), Dynamics of science based innovation.(pp 45–72). Berlin: Springer.

RDISCLOSURE@ database. www.researchdisclosure.com (last visited 28/11/2003).

Salton, G. (1968). Automatic information organisation and retrieval. New York: McGraw—Hill.

Salton, G. (1969). A comparison between manual and automatic indexing methods. American Documentation, 61–71.

Salton, G., McGill, M.J. (1983). Introduction to modern information retrieval. New York: McGraw-Hill.

Salton, G., Wu, H. (1981). A term weighting model based on utility theory. In R.N. Oddy, S.E. Robertson, C.J. van Rijsbergen, R.W Williams (Eds.), Information retrieval research. (pp 9–22). Boston: Butterworths.

Sarasua, L., Corremans, G. (2000). Cross-lingual issues in patent retrieval. ACM SIGIR 2000 Workshops on patent retrieval. Online proceedings. http://research.nii.ac.jp/ntcir/sigir2000ws/ (last visited 28/11/2003).

Schmoch, U. (1997). Indicators and the relations between science and technology. *Scientometrics*, 38 (1), 103–116.

Sparck Jones, K. (1999). *The role of NLP in text retrieval*. In T. Strzalkowski (Ed.), Natural language information retrieval (pp.1–24). Boston (MA): Kluwer.

Swanson, D.R. (1986). Fish oil, Raynaud's syndrome, and undiscovered public knowledge. *Perspectives in biology and medicine*, 30 (1), 7–18.

Tijssen, R.J.W., Korevaar, J.C. (1997). Unravelling the cognitive and interorganisational structure of public/private R&D networks: A case study of catalysis research in the Netherlands. *Research Policy*, 25, 1277–1293.

Tijssen, R.J.W (2004). *Measuring science—technology interactions*. In W. Glaenzel, H. Moed, U. Schmoch (Eds.), Handbook of quantitative science and technology research. Kluwer Academic Publishers.

Turner, W.A., Buffet P., Laville F. (1991). LEXITRAN for an easier public access to patent databases. *World Patent Information*, 13 (2), 81–90.

Verbeek, A., Debackere, K., Luwel, M., Andries, P., Zimmermann, E., Deleus, F.(2002). Linking science to technology: Using bibliographic references in patents to build linkage schemes. *Scientometrics*, 54 (3), 399–420.

Weeber, M., Klein, H., Aronson, A.R., Mork, J.G., De Jong-Van Den Berg, L.T.W., Vos, R. (2000). Text based discovery in biomedicine: the architecture of the DAD–system. *Journal of the American Medical Informatics Association*, Suppl., 903–907.

Wong, S.K.M., Yao, Y.Y. (1995). On modelling information retrieval with probabilistic inference. *ACM transactions in Information Systems*, 13 (1), 38–68.

Zipf, G. (1949). *Human behaviour and the principle of least effort*. Reading (MA): Addison-Wesley.

Zitt, M., Ramanana-Rahary, S., Bassecoulard, E. (2003). Bridging citation and reference distribution.: 1 — The referencing structure function and its application to co-citation and co-item studies. *Scientometrics,* 57 (1), 93–118.

Chapter 31

MEASURING AND EVALUATING SCIENCE–TECHNOLOGY CONNECTIONS AND INTERACTIONS

Towards International Statistics

Robert J.W. Tijssen

Centre for Science and Technology Studies (CWTS), Leiden University, the Netherlands
E-mail: tijssen@cwts.leidenuniv.nl

Abstract: Abstract: Despite the generally acknowledged importance of science in many high-tech areas of major economic relevance, there are no science–related statistics to be found in high–profile international benchmarking reports such as the European Innovation Scoreboard. Why? This chapter aims to provide an answer by advancing our understanding of the possibilities of indicators quantifying linkages between science and technology. Central are the concepts of 'innovation capability' and 'science/technology interface', which are used to assemble a wide range of empirical studies and quantitative indictors to summarise their possibilities and limitations for producing comparative statistics. The review focuses on indicators dealing with flows of written ('codified') information, and indicators of inventiveness that capture the non-codifiable 'tacit knowledge' dimension. General conclusions will be drawn with a view towards further developments in the foreseeable future, suggesting new avenues for the design and implementation of patent–based and inventor–based statistics to describe and assess the complex and dynamic web of relationships between scientific research and technical development within the context of regional or national systems of innovation.

1. SCIENCE/TECHNOLOGY INTERFACES AND INNOVATION CAPABILITY

Technical change and technological innovation have become major drivers of economic progress in the knowledge oriented capitalist economies where growth, productivity, and competitiveness are increasingly based on

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 695-715.

improved technologies, novel products, upgraded processes or customized services. The creation of radically new knowledge, modifying or improving existent knowledge, or imitation of others, has become central to economic development. New discoveries, state–of–the–art information gathering procedures, or successful problem solving routines are often at the core of these innovations. Suffice to say that innovative activity relies heavily on knowledge creation processes and 'intangible assets' such as creativity, know–how, brain power and human experience that have thus become the most valuable resources of our time, much as raw materials were during the early times of industrialization. Even in those days the utilitarian value of the economic asset 'scientific knowledge' was clearly acknowledged: "The value and even the mark of true science consists in my opinion in the useful inventions which can be derived from it", according to the famous German scholar G.W. Leibniz (1646–1716). Scientific research and engineering may indeed have a large impact on technical development and inventions, a fact borne out by several empirical studies (e.g., Salter and Martin, 2001; Cohen et al., 2002). The effectiveness of knowledge creation and knowledge flows between individuals and organizations that are active in research and development (R&D) has become a very important competitive factor in advanced economies. Highly qualified R&D staff, especially those active at the frontiers of scientific research in strategic fields of economic relevance, is considered to be one of the key sources of future prosperity.

However, the effect of 'up stream' R&D on 'down stream' technical inventions and related technological innovations tends to be time delayed, indirect and partial (e.g., Adams, 1990). Moreover, providing direct contributions to achieving higher levels of productivity and economic welfare is not considered to be one of the main missions of public research organizations (PROs) that are active in knowledge production, especially those conducting basic research, although two fairly recent sociological theories of science emphasize that modern science is increasingly inclined to connect with other sectors of society in ways that signal a structural changing towards meeting economic needs. In doing so, new ways of doing science, new partnerships, responsibilities, and more varied funding arrangements have gradually emerged in the 1980s and 1990s. Gibbons et al. (1994) introduced the term 'Mode 2' science for this new type of scientific activity that, amongst other things, emphasizes research teams (rather than individual scholarship), interdisciplinary research (rather than mono-disciplinary), and closer partnerships with industry (and less 'ivory tower' research disconnected from societal needs and business interests). According to the Mode 2 model, science is gradually progressing towards a technological orientation rather than a theoretical orientation. The Triple Helix theory launched and extended by Etzkowitz and Leydesdorff (2000) develops along the same lines, but stressing the role of government along with the other two helices: universities and industry.

Knowledge creation and utilization processes are complex, dynamic, and non-linear. For simplicity's sake we discern a linear sequence of five stages: creation; dissemination; acquisition; storage; and absorption of knowledge . Although (intermediate) results of this process will materialize in written form ('codified'), the entire process rests mainly on human actions and interactions generating 'tacit' (non-codifiable) knowledge. Effective communication and exchange of ideas, results, and experiences often occur through close personal contacts and professional networks. Hence knowledge stocks and flows within R&D and innovation processes are not only critically dependent on the supply and demand for state of the art knowledge, but also on cognitive skills and the ability to adapt to continuous change through lifelong learning. Both PROs and R&D intensive technology firms alike create scientific and technical knowledge through this dynamic process of information gathering and the transformation into person embodied (tacit) know-how, expertise, and skills. Individual researchers, engineers, and technicians are primary actors in these learning and knowledge creation processes, embodying the local capabilities that may be combined at the organizational level (Scott and Bruce, 1994; Mumford and Simonton, 1997).

Nonaka and Takeuchi's (1995) model of organizational knowledge creation and communication captures key aspects of these processes by introducing the following four concepts: (1) 'socialization' that takes place in the case of tacit–to–tacit conversion, when experiences are shared through training or by conversation; (2) 'combination' refers to explicit–to–explicit conversion that occurs when an individual or group combines discrete elements of explicit information into a new piece of codified knowledge, thereby adding value; (3) 'externalisation' is a process in which tacit–to–explicit conversion occurs when an individual or group is able to articulate the foundations or key elements of tacit knowledge; (4) 'internalisation' happens when new explicit information is shared through tacit–to–explicit conversion processes. Information is absorbed and transferred amongst organizational members, and knowledge is developed until new scientific or technical knowledge leads to technical inventions and downstream innovations.

This conceptual framework can be extended to extra-organizational interactions within professional networks or within a 'techno–science' community at large (Rappa and Debackere, 1992). Following the Nonaka/Takeuchi model one can view this process as one in which 'organizational sets and routines' (Nelson and Winter, 1982), turn originally unstructured data into unique tacit or codified knowledge that is acquired by individual researchers and/or inventors, or shared knowledge amongst members of the same organization or network, or disseminated throughout a wider community. As more and more R&D workers become involved in these information conversion and dissemination processes, the more the associated R&D activities and knowledge generating processes become productive and effective in creating societal relevant added value. The 'communication space' in which all these interactions and transitions take place can be coined the 'science/technology interface (which will be referred to as the 'interface' from on now). Obviously the communication processes and feedback loops within this interface are complex, interrelated, and multi-directional. Moreover, the objects, such as research reports or technical manuals, and the subjects (PhD students, researchers, technicians, inventors etc.) related to knowledge flows are varied, interrelated, and may take on different roles at different stages in communication processes. The outcomes of these processes may, for example, take on the form of technical artefacts that have (potential) commercial value with patenting activity at the end of the R&D pipeline. In other cases, where research–based technical knowledge is not appropriated or appropriable, the (intermediate) key results might be published in the open scientific and technical literature.

Many new and interesting developments in the interface are now emerging and flourishing at the cross roads of different knowledge scientific disciplines and technical areas, such as nanotechnology, often driven by pooling of R&D expertise and joining forces across organisational, sectoral or national boundaries. As for the latter, according to recent statistical analysis of worldwide trends in public–private co-authored scientific publications, notably the research articles published in international scientific and technical journals that were jointly authored by industrial researchers and academics, a significant growth in cooperation did indeed occur in the mid and late 1990s. This pervasive trend reflects the increasing orientation of academics towards industrial relevant research and joint knowledge creation with partners within the private sector (Tijssen, 2004a).

Today many regions and nations actively promote R&D cooperation and knowledge transfer specifically aiming at strengthening their 'innovative capability' which would then typically lead to the introduction of novel product oriented or process oriented technologies, or even the start of new R&D–based technology companies. Figure 31.1 depicts the main determinants of innovation capability, as well as their overlaps and interrelations in terms of affecting internal (knowledge–related) factors and external ('framework') factors contributing to R&D processes that shape technological and innovation potential. The broad and somewhat fuzzy concept of innovative capability can provide important insights into the technological leadership and innovative potential of specific firms, industries, regions or countries. Measuring the performance of these entities for each of those determinants provides a valuable indication of competitive strengths and weaknesses and how they stand as potential sources of new technologies. Improvement of their innovative capability of firms may depend on allocating more resources for R&D, promoting first–mover research, engaging in research alliances with other firms, or outsourcing non-research activities, while public domain policies can be devised for improving framework conditions and enhance human competences by supporting R&D programmes and technology transfer mechanisms, and upgrading of education and training facilities.

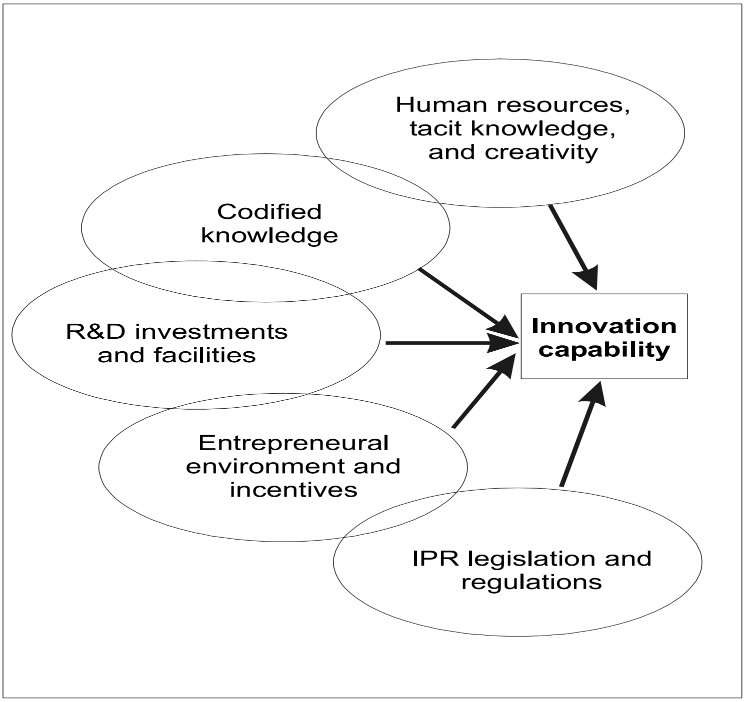


Figure 31.1. Diagram of determinants of innovation capability (source: Tijssen, 2003b)

Adopting this concept is also beneficial for quantitative modelling and developing diagnostic indicators to register and monitor changes in the role and contributions of both codified knowledge as well as 'tacit' human resources in building and maintaining R&D–based innovation potential within innovation systems. Both nodes and flows are important in these systems, since knowledge diffusion and spill-over processes, combined with absorptive and learning capacities among agents ('actors') in the system, determines the system's distributive power and its effectiveness. This 'systems view' has further implications for comprehensive statistical analyses, since science–based knowledge production, and utilization of that knowledge for technological development, is important for the performance of the entire system irrespective of the location of the R&D performing agents — either public research or in the private sector.

2. INDICATORS AND STATISTICS

With the advancement of globalisation and the drive towards knowledge oriented economies it has become more important than ever to establish comparative benchmarks for the performance of national and regional innovation systems. In recent years, innovation indicators of have come into regular use, and internationally comparative innovation statistics are published widely by statistical agencies and (supra)national governments. The European Commission's European Innovation Scoreboard contains a wide range of national statistics, including some dealing with R&D expenditures, counts of (high tech) patents, and quantities of science and engineering graduates (EC, 2003b). However, data on the contribution of science — neither basic science nor applied science — to technological development is conspicuously absent. At the same time the economic and social dimensions of science/technology interactions are becoming more important than ever before for innovation–related policy analysis and decision–making at both regional and national levels. As a consequence, an increased interest has arisen for evidence–based assessment and statistics on the science/technology interaction. This need for 'hard' comparative numerical data is yet to be fulfilled. One of the main reasons is that measuring contributions of science to technological innovation, and creating databases and appropriate quantitative indicators to do so, are still subject of research and scholarly debate. Despite the theorizing, and the empirical case studies have been conducted during the previous two decades that were aimed at grasping key characteristics of processes by which scientific and technical knowledge actually drives technological advances, the impact of R&D on innovative products and processes is one of the most difficult areas of scientific investigation to understand and model in terms of quantitative measurement and statistics (e.g., Coombs and Hull, 1998; Hagedoorn et al., 2000; Salter and Martin, 2001; Cohen et al., 2002). The importance of governance systems, organisational structures and individual factors has long been recognized, but their contributions in enabling, shaping and driving knowledge creation, transfer and utilization processes remains complex and eludes systematic large-scale comparative analysis.

Table 31.1. Classification of knowledge transfer/exchange mechanisms

Mechanism	*Knowledge type*	*Measurability*	*Data availability**
Informal contacts, networks	tacit	=	=
Shared technical facilities	tacit	+	=
Formal contacts, networks	tacit	+	+
Education and training	tacit/codified	+	+
Contract research, consultancy	tacit/codified	++	+
Public/private R&D cooperation	tacit/codified	++	++
Human resources and mobility	tacit	++	++
Spin-off/start-up companies	tacit	+++	++
Patents (citations)	codified	+++	+++
Research papers (citations)	codified	+++	+++

* Availability of internationally comparable statistics, (or comprehensive information sources that enable the production of those statistics). Indication of measurability/data availability:

+++ Good measurability/internationally standardized statistics;

++ Reasonably good measurability/some international statistics;

+ Limited measurability/national measurements or statistics only;

= Only qualitative data /(inter)national statistics not available.

The host of empirical studies have succeeded in unearthing a multitude of intertwined processes and mechanisms by which scientific and technical knowledge is created and disseminated across the interface to a wide range of users (e.g., OECD, 2002a). Table 31.1 provides a non-exhaustive list of major knowledge transfer mechanisms facilitating knowledge related exchanges and interactions between public research sector and corporate R&D. Attached to this shortlist is a crude typology of their potential as information sources for international statistics. Each mechanism is characterised in terms of the primary type of R&D related knowledge it represents, a general assessment of its measurability for cross-country comparisons, and the current availability of internationally comparable statistics. Even though some mechanisms and associated knowledge flows are inherently difficult to measure within a comparative benchmarking framework (notably, informal contacts), several of the others are amenable to systemic quantification for statistical applications. However, at present only patents and research publications (referred to as a 'research paper' from here on) enable reasonably standardized quantitative information enabling both large-scale systemic and in-depth international comparisons across different countries, knowledge domains and institutional sectors. Patents and research papers contain a rich variety of relevant information on both the objects and subjects of knowledge flows. Imagine the following extreme example: a corporate owned biotechnology patent citing a large number of papers publishing scientific breakthroughs or novel techniques in a range of life science areas, listing several co-inventors employed by public sector research organizations and universities, of which some co-inventors are also authors of research papers cited in that patent, and other co-inventors also have now left the public sector to work in corporate R&D or have (helped) launch spin-off companies.

The following sections will elaborate on statistical methods and quantitative indictors to characterize further the information potential of research papers and patents as information sources for producing international statistics on interactions and knowledge flows within the interface. Where research papers represent scientific progress, especially basic scientific research, patents represent technical developments and are a good candidate to shed further light on the R&D–based innovation potential in certain technical areas. Using in the bibliographic data contained in both document types one can discern the following categories for measuring and evaluating various aspects of science/technology linkages:

a) Corporate research papers, i.e. produced by R&D staff in the private sector;
b) Public/private co-authored research papers;
c) Citations in corporate research papers to public research papers;
d) Citations in patents to research papers;
e) Patents (co-)produced by research staff at PROs;
f) Inventors who also publish research papers.

Section 3 deals specifically with R&D–based knowledge flows as reflected by patent→paper citation indicators, reflecting Nonaka/Takeuchi's combination processes within the interface. Section 4 focuses on the role of (co)inventors, thereby emphasizing the role of socialization and externalisation processes. The latter 'human factors' dimension of science/technology connectedness obviously also touches on other policy–relevant issues of human resources; more specifically the mobility of PhDs and skilled R&D workers from the public research sector to industry, which is now considered a high–profile science/technology knowledge flow and in dire need of internationally comparable statistical information to assess the direction and magnitude of brain drains at regional, national and global levels . The categories (a) to (c) refer primarily to the institutional dimension of science (i.e., science/industry relationships or university/industry relationships), rather than reflecting science/technology interactions, and are therefore excluded from this review.

3. MEASURING KNOWLEDGE FLOWS

The reference lists which are added to research papers and to patents contain documentary sources deemed relevant for substantiating and delineating intellectual property claims, or to describe the scientific knowledge base and preceding technologies (the 'prior art') in the field. As such, these references reflect linkages between the patent and 'cited' source offering detailed empirical information about the dissemination and absorption of that codified knowledge. This knowledge flow approach of communication processes encompasses a variety of 'citation' indicators in which one document refers to another: (a) paper→paper; (b) patent→patent; (c) patent →paper, and sometimes, although still less frequent; also (d) paper→patent citations (e.g., Hicks, 2000). Analysis of all the citation traffic within the interface, i.e., categories (c) and (d), provides quantitative data on knowledge flows at the interface in terms of their sources, direction, intensity, and users of knowledge flows within and across geographically, institutionally and disciplinarily defined borders.

So far, most of these citation flow studies are restricted to the analysis of citation links between documents stored in three large bibliographic databases: (1) the Thomson/ISI databases, especially the *Science Citation Index®* (SCI) or the *Web of Science®* (its web–based version), with their comprehensive coverage of research articles published in many thousands of peer–reviewed international scientific and technical journals; (2) the patent database of the *European Patent Office* — EPO; (3) the patent database of the United States Patent and Trademark Office — USPTO. These two patent databases have been key sources of statistical information for economic studies of technical change and technological innovation since the 1970s even though a number of important caveats apply to these patent data . Nonetheless, patents provide a detailed and verified source of comparative empirical information on inventive activity, as well as, under specific conditions, offering the added bonus of enabling more detailed analyses of R&D processes .

Patent applicants or patent examiners citing the relevant prior art, with a direct or indirect bearing on the knowledge claims stated in the application, often include one or more research articles in scientific and technical journals that contributed materially to the product or process to be patented. The statistical analysis of patent→paper citation links was pioneered by Francis Narin and his co-workers, using the references within USPTO patents to the research articles published in international scientific and technical journals (e.g. Carpenter and Narin, 1983; Narin and Noma, 1985). Since those early days quite a substantial literature has evolved on patent–based empirical studies of the complex web of interrelationships and knowledge flows between published science and patented technology. We now have ample empirical evidence about which specific fields of science are of relevance for technical development in 'science based' fields of technology. Most of these studies were done at the aggregate level, focused on general features of these linkages, and produced some statistics at the level of fields of science or technology areas (e.g., Brusconi et al., 2003), or even at the macro–scale level of entire countries (e.g., Grupp and Schmoch, 1992; Narin et al., 1997; Hicks et al., 2000; Tijssen, 2001; Verbeek et al., 2003). Patent→paper citation statistics now feature quite prominently in both the National Science Foundation's Science and Engineering Indicators Report (NSF, 2002) as well as the European Commission's DG Research Third European Report on Science and Technology Indicators (EC, 2003a), which released statistical tables with citation frequency data as proxy measures of science/technology linkages within broad fields of science.

Parallel to these large-scale citation analyses, several small-scale case studies were carried out aimed primarily at validation and contextualization of findings emerging from the either patent→paper or paper→patent citation studies (Carpenter and Narin, 1983; Van Vianen et al. 1990; Albert et al., 1991; Schmoch, 1993; Meyer, 2000a; Tijssen, 2000). Collectively this body of evidence confirms that reference lists in patents represent valuable information on explicit connections between scientific and technical knowledge and technical inventions; information which can be used for describing systemic features of the interface. However, these studies also clearly indicate that the connections reflected by citations are not necessarily causal links: in other words, patent→paper citation data are more appropriate for statistics on the interaction between science and technology, rather than the strength of those linkages or the degree of connectedness. Furthermore, Narin et al. (1997) revealed the existence of domestic self-citation propensities in all major countries — i.e., a relatively large share of the citations, in the range of two to four times more than statistically expected, refer to research papers originating from the same country. This so-called 'domestic bias' in patent citation relations of this magnitude clearly indicates localized knowledge flows, suggesting relatively strong interactions between scientific and technological progress as well as cumulative effects in knowledge creation and dissemination in regional or national R&D systems and innovation systems (e.g., Hicks et al., 2001). For example, in the case of the Netherlands, Dutch invented USPTO patents cite Dutch scientific papers four times more often than expected after controlling for the size of Dutch science (Tijssen, 2001). The majority of these domestic citations proved to be author/inventor self-citations, i.e., inventors citing their own research papers. These self-citation links reflect direct — if not causal — relationships between research and technological development, which can be used as a valid indicator of person embodied science/technology linkages. It seems more than likely that the domestic citation patterns of other countries are also significantly affected by the same level of self-citation propensity, especially in domains of the interface where (basic) scientific research and technology development are closely connected such as in the case of biotechnology.

Thus far most citation studies have focused almost entirely on citations to research articles in ISI covered journals that are listed on the front page of USPTO patents. Since EPO examiners tend to focus much more on the patent literature than the non-patent literature when reviewing the prior art, the references listed in EPO examiner's search reports are less appropriate for comprehensive and comparative patent→paper citation analyses (Michel and Bettels, 2001). Interestingly, recent citation impact statistics published by the European Commission (EC, 2003a) rely entirely on EPO patents at the danger of significantly misrepresenting both the magnitude and dynamics in the science relatedness of technical areas.

In summary, the current statistics on citations between patents and research papers in international journals appear useful but are nonetheless a rather crude reflection of science/technology knowledge flows. In the light of these drawbacks further improvements should tackle the following three methodological issues at the very least to render patent→paper citation statistics more amenable for valid international comparisons:

1. Self-citation propensities. A clear cut breakdown by geographical proximity should be made between 'local' author/inventor self-citations, domestic citations, and foreign citations. Each citation flow is likely to be driven by different knowledge conversion processes and different communication channels, with varying degrees of relevance in terms of direct linkages between patented technologies and cited science;
2. Relevance of EPO patents. Further research is needed as to the degree of relevance of EPO patents compared to USPTO patents, especially for those interfaces where patent →paper citations are scarce and therefore probably too low for meaningful statistics. One of the benefits of the EPO system, however, is its labelling of each citations in terms of relevance for the patent claim, information that is now also available in Questel/Orbit's on-line version of the EPO patent database;
3. Non-journal patent→paper citations. It is not yet clear to what extent the documents, other than research papers in scientific or technical journals, which are cited in patents actually reflect contributions from basic or applied research — either from PROs or from corporate laboratories. Further case studies and macro-scale comparative analysis of these references may well prove to be a goldmine of relevant information on

contributions of applied research and technical development. Harvesting this source will require sophisticated text processing routines and algorithms (Lawson et al., 1996).

Citation traffic between documents is by definition a two-way street: the research papers in the scientific and technical literature can obviously also cite patents. Recent empirical studies indicate that these paper→patent citations tend to be a rarity: only 1% of the research articles in the SCI database cite USPTO patents (Glänzel and Meyer, 2003). It would seem that such reverse flows are most likely to occur in the highly interactive interfaces characterized by many feedback loops between scientific research and technical development, where patents also tend to cite the research literature. The meaning of such citations is still a subject of research: are technological developments leading scientific progress in these cases? Or are there 'human factors' at play, such as a larger propensity for reverse citing by prolific inventor–researchers working at corporate R&D labs of technology firms and citing their own patents or other in-house patents?

4. MEASURING HUMAN FACTORS

Obviously neither research methods, nor tacit knowledge, nor technical artefacts can be communicated and transferred in full through research papers and patents. Naturally citation flows cannot capture the essence and context of knowledge conversion processes and knowledge utilization mechanisms. However, the citation flows can be used to pinpoint important agents and actors in these transfer and utilization processes of codified knowledge. In this section will tap into those communication processes and focus on the inventor — i.e., the researcher, engineer, or technician responsible for initial idea, the watershed discovery or major technical breakthrough leading to that invention. Inventor–based analysis enables us to go inside the 'black box' of person–embodied tacit information and to quantify characteristics of human capital and intellectual capital underpinning knowledge creation and transfer at the interface. Especially in those cases where these individuals also participate in follow-up stages (R&D, patenting and commercialisation), they often hold the key for unraveling important details and understanding process characteristics of the entire innovation trajectory (Tijssen, 2002).

Invention is a complex interactive path driven by intellectual effort, inspiration, creativity, incentives and reward systems. Technological inventiveness seems to concentrate within a precious few highly talented and competent people. Using patent information, Narin and Breitzman (1995) and Ernst et al. (2000) found high concentrations of patents in very small numbers of inventors who appear to be largely responsible for innovations and competitive success of technology companies. Most likely these individuals are internally driven to push back frontiers — whether they are motivated by curiosity, the need for personal achievement, money, power or fame, they pursue knowledge creation and application systematically. They also actively seek improvements and search for new applications — the root of all innovation — with a eye for the potential of those novelties in terms of financial rewards and economic return on investment. Invention often involves a leap unto the unknown, where trial and error, the unexpected or even chance can have a significant impact on the inventive trajectory and its final outcome.

How successfully inventors will internalise and translate transferred scientific information and research knowledge into their own inventiveness will depend largely on their absorptive capacity, and their ability and commitment to learn and create new knowledge. This obviously applies to the group of academic inventors who work at PROs and are most likely to bridge successfully the domains of science and technology. These inventors need to stay abreast of the newest developments in relevant research, and may even play a leading role in scientific progress. The corporate inventors, especially those working in central research labs are often specifically hired to invent. A third group of inventors are the ‘independents’, i.e. those without corporate or other institutional affiliation. This very heterogeneous group consists of entrepreneurial inventors who may have launched their own spin-off or start-up companies to invent or commercialise their invention, inventors who licensed their inventions, but also those who have (so far) failed to exploit their patent(s).

Focussing the inquiry on the names of inventors listed in science related patents, rather than analyzing the reference list of such patents, opens up a second possibility for tracing (person embodied) knowledge flows. It also enables an in-depth systematic analysis of cooperation patterns and network linkages between people listed as inventors. Pursuing the latter avenue, a new string of patent–based empirical studies is gradually emerging aimed at tracing R&D staff who are simultaneously active in ‘R’ and ‘D’ (e.g., Noyons et al., 1994; Meyer, 2000b; Tijssen, 2002). Of particular interest are, of course, those individuals who visibly bridge both knowledge domains: researchers publishing in the scientific literature and those producing patented inventions. This may refer to one and the same person, or encompass several people linked by joint publications and/or joint patents. The study by Balconi et al. (2002), dealing with linkages between Italian inventors and Italian university researchers, is perhaps one of the first to use this human resources oriented approach for mapping structural characteristics of the interface at the level of an entire national innovation system. One of the key methodological issues in their study is the degree to which the organizational environment affects linkage patterns and knowledge flows, where Italian academic inventors are more connected than their non-academic counterparts. A case study amongst Dutch inventors confirms that organizational factors are of pivotal importance for understanding and modelling the 'science dependence' of patented inventions; not only should the 'cognitive distance' between scientific and technological areas be taken into account in these inventor studies, but also the R&D mission of institutions and companies, as well the R&D environment in which inventive activities take place (Tijssen, 2002, 2003a).

Another fruitful entry point for tracking explicit links between scientific research and technological developments is to examine the patents in research intensive fields of technology in which one or more staff members of PROs lay hidden amongst the list of inventors. Because PROs often leave ownership of the patent to the firm(s) which financed the research project, the contribution of these PROs remains virtually invisible owing to lack of information about inventor affiliations on the front page of the patent. These co-inventors of corporate–owned patents represent direct relationships, and most likely quite strong connections, between academia and industry, especially in the science–based industrial sectors such as biotechnology, pharmaceuticals, and medical instruments. Studies by Balconi et al. (2002) and Meyer (2003), but also by Saragossi and Van Pottelsberghe de la Potterie (2003), provide clear empirical evidence that the number of university invented patents is much higher than the number of university–owned patents; in fact, the number of these EPO patents for the Université Libre de Bruxelles is more than double the number of university–owned patents for the whole period 1985–1997. Balconi et al. (2002) identified that out of 1,300 university–invented patents in Italy during the years 1979–1999 only 90 EPO patents had university assignees, whereas Italian university–invented patents account for 3.8% of EPO patents by Italian inventors. Meyer (2003) reports that Finnish universities owned 36 USPTO patents in the period 1986–2000, but that there were 530 Finnish university–invented patents. Germany shows a similar pattern: university–owned patents are relatively rare, but university–invented patents have continuously increased from less than 200 in the early 1970s to around 1,800 in the year 2000 (OECD, 2002b).

Identifying university co-inventors seems therefore a particularly fruitful approach for uncovering and evaluating direct linkages between academic research and corporate applications of those research findings. Since most patents usually list only the inventor's country of residence or the private address, the key methodology issue is: how can one identify and track down those co-inventors? An EC–funded European–wide study, which was carried out by Noyons et al. (2003), extracted information on the inventor addresses from EPO patent files in the life sciences and nanoscience. Their main methodological conclusions were: (1) most addresses listed on the EPO patents are indeed private addresses and therefore not very useful as the sole information source for systematic exploitation; (2) matching inventor names with author names of research papers covered by the *Science Citation Index* (SCI) provides an additional, and more fruitful, route for gathering the institutional addresses of those author/inventors, especially those who are active in science–based technical areas.

Findings from another recent pilot study, conducted in the Netherlands amongst university co-inventors on the basis of USPTO and EPO patents filed by companies and published in 2002–2003, and using the SCI as one of two sources to track down those co-inventors (Tijssen, 2004b), provides further empirical evidence of the contributions of public sector science to private sector technical development: 59% of these university co-inventors judged their research contribution to be of crucial importance in the R&D leading to the patent; more importantly, almost 80% of the inventors that were interviewed indicated that the list of inventors on their patent included all academics with key contributions to the inventive process.

The inventor names listed on patents opens up the possibility for developing thesauri and typologies of inventors (e.g., university co-inventors). Moreover, linking inventor names to their previous and current affiliate addresses enables statistical analyses of inventor mobility between employers, institutional sectors and countries (Tijssen, 2003b). A very interesting recent development is the PatVal project, an on-going EC funded study based on a number of large-scale surveys in selected EU-15 member states . PatVal collects a diverse range of data on EPO patents filed during the years 1992–1997, including estimates of their perceived monetary value, but also background information about the inventors and their organizational environment and employer mobility. For example, the first results of the PatVal survey conducted in the Netherlands shows that 30% of the inventors have moved to a different employer since the patent was filed (Verspagen, 2004).

Given the results presented above, it seems fair to conclude that important first steps have now been taken towards developing statistically robust methodologies and constructing comprehensive databases for international comparisons of science/technology connections. Inventor related statistics and university–invented/corporate–owned patents are now also on the research agenda of international agencies producing statistics such as the OECD, but it is still early days in terms of comparative measurements and indicators with the potential to produce widely accepted statistics. Further studies are required to describe and model key characteristics of the interface between scientific research and technological development. These studies should focus on the R&D workers — not only the prolific 'star' inventors in the corporate sector, or the author/inventors employed by PROs, but also the many independent inventors who are likely to represent the backbone of innovation capability in less developed countries.

5. THE WAY FORWARD

It is an encouraging sign that patent citations statistics are now being used for internationally comparative objectives (NSF, 2002; EC, 2003). This patent→paper citation approach is certainly an interesting way of looking at knowledge flow patterns within science/technology interfaces, but it is still in its infancy in terms of valid aggregate–level citation statistics. However, the strengths and limits of its descriptive and analytical power, and its potential for reliable cross-country comparative statistics, are yet to be fully determined. If and when patent citations are to be used for modelling nation–specific characteristics of science bases, or for comparative measurements of the importance of science/technology linkages within national innovation systems, it would at the very least seem wise to incorporate a base of patents as wide as possible (e.g., the so-called 'triad patents' covering EPO, UPSTO and the Japanese Patent Office JPO), whilst making a clear distinction between international citations, national citations and author/inventor self-citations. The internationally comparable statistics that are currently published by the OECD, on the basis of triad patents, is certainly a step in the right direction (OECD, 2003). At present the scope for policy–relevant statistics would seem to be restricted to those industrial sectors and technical areas in which patents are one of the essential vehicles for protecting intellectual property rights, but of those fields of science where research papers in journal are valid representatives of the contributing scientific knowledge.

Whilst the development of indicators of knowledge creation and knowledge flows now enjoys a relatively high status in academic research, and has attracted policy attention and funding, there is much less interest in indicators of science–related innovation capability. This is surprising in view of the pivotal role of 'human factor' in the early stages of innovation process and the critical importance of the existent stock of intellectual capital and human capital for sustained economic development. Statistics on the distribution of different kinds of inventors, characteristics of their inventive performance, and their institutional locations and working environments, could add very valuable insights and data to help quantity and understand the inner workings of science–based R&D processes within innovation systems. To capture this source we need better models, and new methodologies and metrics, which deal explicitly with characteristics of scientists and inventors who produce (patented) technologies. Case studies and national surveys can contribute to a better process understanding of the role of R&D staff and inventors working in the science/technology interface, especially within industrial–relevant fields of the medical sciences and life sciences (e.g., Malo and Geuna, 2000; McMillan et al., 2000).

Clearly, both classes of patent–based measurements (i.e., patent citation related and inventor related) represent at best useful proxies of science/technology interactions and human creativity relevant to industrial innovation. In both cases they fall short in terms of their potential to be singled out as widely accepted leading indicators of R&D related aspects of innovation capability. Given the current stage of development of measurement techniques, databases, and survey methodologies the scope of utilization lies mainly within their ability to analyse and compare knowledge creation and dissemination characteristics, rather than the direct contribution of R&D outputs in innovation trajectories. The quantum leap to truly innovation related statistics requires additional sources of comparative data, as well as data collection methodologies aimed at achieving internationally standardized ('harmonized') data — a step to be taken not from a naïve perspective focussing on measurement methodologies solely. We also need a contextualised assessment of all data at hand using sound theory and appropriate conceptual models. Although many recent developments in theories and sophisticated generic models seem to have outrun the ability of the available statistical material to provide verifiable factual evidence (e.g., Bozeman 2000; Schmoch et al., 2000, Joanneum Research et al., 2001; OECD, 2002a), it is now time to use these inputs as a guide to gather systemic empirical evidence of science/technology connectedness and to develop and test new indicators aimed at producing reliable harmonized statistics. In these circumstances it would be well advised to maintain a experimental approach, using multiple databases and methods as a matter of principle, whilst recognizing the need for continuing research in this field.

One of the fundamental issues to be addressed are the intricate relationships between on the one hand country/region specific strengths and critical weaknesses of science bases, in terms of their organizational structures and enabling conditions, and on the other hand the effectiveness of knowledge transfer mechanisms that drive and shape learning processes and knowledge creation processes that eventually lead to (patented) inventions and related innovations. It is obvious that much time and effort needs to be invested and to examine underpinning theoretical notions and hypotheses, to design effective and efficient methodologies, and to develop appropriate indicators. Existing agencies and information collecting infrastructures that are already in place, for instance EUROSTAT and the OECD, may provide technical infrastructures and organisational frameworks for further development and upscaling of data collection routines. Such a dedicated R&D program and concerted effort should help pave the way towards creating information systems that enable large-scale production and analysis of internationally comparative statistics for outlets like the European Innovation Scoreboard (2003b). These efforts should not only help improve the quality and scope of indicators capturing key characteristics of the science/technology interface, or using existing indicators in novel innovative ways, but will most certainly also trigger further promising avenues of research with outcomes that underscore the vital importance of high quality science bases and innovation capability of regions and countries that are progressing towards becoming leading knowledge–based economies.

REFERENCES

Adams, J. (1990). Fundamental stocks of knowledge and productivity growth. *Journal of Political Economy*, 98, 673–702.

Adams, J., Stephan, P., Sumell, A. (2003). *Capturing knowledge: the location decision of new PhDs working in industry.* Presentation at the Roundtable for Engineering Entrepreneurship Research, Atlanta: Georgia Institute of Technology, November 2003.

Albert, M., Avery, D., Narin, F. (1991). Direct validation of citation counts as indicators of industrially important patents. *Research Policy*, 20, 251–259.

Balconi, M., Breschi, S., Lissoni F. (2004). Networks of inventors and the location of university research: an exploration of Italian data. *Research Policy*, 33, 127–145.

Bozeman, B. (2000). Technology transfer and public policy: a review of research and theory. *Research Policy*, 29, 627–655.

Brusconi, S., Criscuolo, P., Geuna, A. (2003). *The knowledge bases of the world's largest pharmaceuticals groups: what do the patent citations to non-patent literature reveal?* SPRU report, University of Sussex, United Kingdom.

Carpenter, M., Narin, F. (1983). Validation study: patent citations as indicators of science and foreign dependence. *World Patent Information,* 5, 180–185.

Cohen, W., Nelson, R., Walsh, J. (2002). Links and impacts: the influence of public research on industrial R&D. *Management Science,* 48, 1–23.

Coombs, R., Hull, R. (1998). Knowledge management practices and path dependency in innovation. *Research Policy*, 27, 237–253.

EC (2003a). *Third European Science and Technology Indicators Report*. Brussel: European Commission.

EC (2003b). *European Innovation Scoreboard 2003*. SEC(2003) 1255, Brussel: European Commission.

Ernst H., Leptien, C., Vitt J. (2000). Inventors are not alike: the distribution of patenting output among industrial R&D personnel. *IEEE Transactions on Engineering Management*, 47, 184–199.

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.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M. (1994). *The new production of knowledge: the dynamics of science and research in contemporary societies.* Sage, London.

Glänzel, W., Meyer, M. (2003). Patents cited in the scientific literature: an exploratory study of 'reverse' citation relations. *Scientometrics*, 58, 415–428.

Grupp, H., Schmoch, U. (1992). *Perceptions of scientification of innovation as measured by referring between patents and papers: dynamics in science–based fields of technology*. In: Grupp, H. (Ed.), Dynamics of science–based innovation. Berlin: Springer–Verlag.

Hicks, D. (2000). 360 Degree linkage analysis. *Research Evaluation*, 9, 133–143.

Hicks, D., Breitzman, T., Olivastro, D., Hamilton, K. (2001), The changing composition of innovative activity in the US — a portrait based on patent analysis. *Research Policy*, 30, 681–703.

Hagedoorn, J., Link, A., Vonortas, N. (2000). Research Partnerships. *Research Policy*, 29, 567–586.

Joanneum Research, in cooperation with ZEW and ARCS (2001). *Benchmarking industry–science relations: the role of framework conditions*. Vienna/Mannheim: Report to the European Commission (DG Enterprise) and the Austrian Federal Ministry of Economy and Labour (www.benchmarking-in-europe.com).

Lawson, M., Kemp, N., Lunch, M., Chowdhury, G. (1996). Automatic extraction of citations from the text of English language patents — an example of template mining. *Journal of Information Science,* 22, 423–436.

Malo, S., Geuna, A. (2000). Science/technology linkages in an emerging research platform: The case of combinatorial chemistry and biology. *Scientometrics*, 47, 303–321.

McMillan, G., Narin, F., Leeds, D. (2000). An analysis of the critical role of pubic science in innovation: the case of biotechnology. *Research Policy*, 29, 1–8.

Meyer, M. (2000a). Does science push technology? Patents citing scientific literature, *Research Policy*, 29, 409–434.

Meyer, M. (2000b). Patent citations in a novel field of technology — what can they tell about interactions between emerging communities of science and technology? *Scientometrics*, 48, 151–178.

Meyer, M. (2003). Academic patents as an indicator of useful research? A new approach to measure academic inventiveness. *Research Evaluation*, 12, 17–27.

Michel, J., Bettels, B. (2001). Patent citation analysis — A closer look at the basic input data from patent search reports. *Scientometrics*, 51, 185–201.

Mumford, M., Simonton, D. (1997). Creativity in the workplace: people problems and structures. *Journal of Creative Behavior*, 31, 1–6.

Narin, F., Noma, E. (1985). Is technology becoming science? *Scientometrics*, 7, 369–381.

Narin, F., Breitzman, A. (1995). Inventive productivity. *Research Policy*, 24 (4), 507–519.

Narin, F., Hamilton, K., Olivastro, D. (1997). The increasing linkage between US technology and public science. *Research Policy*, 26, 317–330.

Nelson, R., Winter, S. (1982). *An evolutionary theory of economic change*. The Belknap Press of Harvard University Press, Cambridge.

Nonaka, I., Takeuchi, H. (1995). *The knowledge creating company: how Japanese companies create the dynamics of innovation*. New York: Oxford University Press.

NOWT (2004). *Science and Technology Indicators 2003 — Summary*. Netherlands Observatory of Science and Technology (NOWT), CWTS/MERIT report to the Netherlands Ministry of Education, Culture and Science (see <www.nowt.nl>).

Noyons, E.C.M, Van Raan, A.F.J., Grupp, H., Schmoch, U. (1994). Exploring the science and technology interface — inventor author relations in laser medicine research. *Research Policy*, 23, 443–457.

Noyons, E.C.M, Buter, R.K., Van Raan, A.F.J., Schmoch, U., Heinze, T., Hinze, S., Rangnow R. (2003). *Mapping excellence in science and technology across Europe — Life Science; — Nanoscience and nanotechnology*. CWTS/FhG–ISI reports for EC/DG Research.

NSF (2002) *Science and Engineering Indicators 2002*. Arlington: National Science Foundation , National Science Board.

OECD (2002a). *Benchmarking Industry–Science Relationships*. Paris: Organisation for Economic Cooperation and Development.

OECD (2002b). *Science, technology and industry outlook*. Paris: Organisation for Economic Cooperation and Development.

OECD (2003). *Science, technology and industry scoreboard 2003—towards a knowledge-based economy*. Paris: Organisation for Economic Cooperation and Development.

Rappa, M., Debackere, K. (1992). Technological communities and the diffusion of knowledge. *R&D Management*, 22, 209–222.

Salter, A., Martin, B. (2001). The economic benefits of publicly funded basic research: a critical review. *Research Policy*, 30, 509–532.

Saragossi, S., Van Pottelsberghe de la Potterie, B. (2003). What patent data reveal about universities: the case of Belgium. *Journal of Technology Transfer*, 18, 47–51.

Schmoch, U. (1993). Tracing the knowledge transfer from science to technology as reflected in patent indicators. *Scientometrics*, 26, 193–211.

Schmoch, U., Licht, G., Reinhard, M. (2000). *Wissens- und Technologietransfer in Deutschland*. Stuttgart: Fraunhofer IRB Verlag.

Scott, S., Bruce, R. (1994). Determinants of innovative behavior: a path model of individual innovation in the workplace. *Academy of Management Journal*, 37, 580–607.

Tijssen, R.J.W., Buter, R.K., 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. (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. (2002). Science dependence of technologies: evidence of inventions and their inventors. *Research Policy*, 31, 509–526.

Tijssen, R.J.W. (2003a). *The knowledge resources of inventions: towards a typology of organizational knowledge creation environments*. Presentation at meeting of the INIR Network, Catholic University of Leuven, Belgium, January 2003.

Tijssen, R.J.W. (2003b). *Inventiveness by numbers: towards inventors statistics*. Invited paper at the WIPO–OECD Workshop on Statistics in the Patent Field, Geneva, Switzerland, September 2003.

Tijssen, R.J.W. (2004a). Is the commercialisation of scientific research affecting the production of public knowledge? Global trends in the output of corporate research articles. *Research Policy* (forthcoming).

Tijssen, R.J.W. (2004b). *De universiteit als verborgen kennisbron: De (on)zichtbaarheid van Nederlandse universitaire co-uitvinders in bedrijfsoctrooien*. CWTS report for the Netherlands Ministry of Education, Culture and Science.

Van Vianen, B.G., Moed, H.F., Van Raan, A.F.J. (1990). An exploration of the science base of recent technology. *Research Policy*, 19, 61–81.

Verbeek, A., Debackere, K., Luwel, M. (2003). Science cited in patents: a geographic 'flow' analysis of bibliographic citation patterns in patents. *Scientometrics*, 58, 241–262.

Verspagen, B. (2004**).** *Rapport over de uitkomsten van het Nederlandse gedeelte van de PatVal enquête onder uitvinders van Europese patented ingediend vanuit Nederland*. Report Eindhoven Centre for Innovation Studies, Technical University Eindhoven.

Chapter 32

THE TECHNOLOGICAL OUTPUT OF SCIENTIFIC INSTITUTIONS

Ulrich Schmoch

Fraunhofer Institute for Systems and Innovation Research, Karlsruhe, Germany; and Institute of Sociology, University of Karlsruhe, Germany
E-mail: u.schmoch@isi.fraunhofer.de

Abstract: Up to now the contribution of scientific institutions to technology is considered to be primarily indirect. However, an analysis of patent applications of European public research institutions shows that they contribute about one half of all patent applications in selected science–based technology fields in the life sciences and nanotechnology. This finding documents a high direct contribution of science institutions to the generation of technology. The share of public non-profit institutions proves to be important, in particular in early stages of the technology life cycle; scientific institutions obviously play the role of lead actors. A comparison between European and German data reveals a lower, but still quite high share of public research institutes in other areas of science–based technology, so that the general statement of a high direct technology contribution holds. With regard to scientific institutions, patent indicators do not replace, but rather complement publication indicators and reflect an additional dimension of performance.

1. INTRODUCTION

Until the middle of the twentieth century the division of labour between universities and other public research organisations on the one hand and industrial enterprises on the other hand was clear-cut. Public research institutions produced pure, basic knowledge and enterprises technological, applied knowledge. With the rise of knowledge–based technologies, impressively described by Freeman (1982), the enterprises engaged in research to a substantial extent and the public research organisations were

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 717-731.

increasingly active in "finalised research" (Böhme et al., 1978) which according to the present wording is called "(application–)oriented basic research" (OECD, 1994, p. 69). Some authors already state a de-differentiation between scientific institutions and industry (for instance Nowotny et al., 2001, p. 21 ff.), but the empirical evidence for this thesis is quite weak. Nevertheless, a growing orientation of public research institutions, in particular universities, on application is reflected in indicators such as external funding by industry or patenting of public research institutes. Although in recent years the role of public scientific institutions in technology is emphasised, most authors primarily see an indirect contribution to the generation of new technology. In a review paper Salter and Martin (2001) analysed the impact of basic research on economic advance and found as major transfer mechanisms:

– "increasing the stock of useful knowledge;
– training skilled graduates;
– creating new scientific instrumentation and methodologies;
– forming networks and stimulating social interaction;
– increasing the capacity for scientific and technological problem–solving;
– creating new firms" (Salter and Martin, 2001, p. 520)

I myself found by a survey at German university institutes that major transfer mechanisms are:

– co-operative research of universities and industrial laboratories;
– informal contacts such as telephone calls or meetings; or
– education of graduates.

Direct mechanisms, especially contract research, appeared to be less prominent in most disciplines, with the major exception of mechanical engineering (Meyer-Krahmer and Schmoch, 1998). The assumption of a limited direct contribution to technology is supported by the finding for various countries that the share of patents of public institutions within all domestic patents is in the range of five to eight per cent (Meyer, 2003; Schmoch, 2000, p. 24 ff.). Thus their contribution to technology is visible, but not overwhelming. However, it can be assumed that the direct contribution of public institutions to technology is more relevant in science–based fields of technology which should be reflected in higher patent activities of public institutions. In this article, I shall analyse this assumption in more detail.

2. APPROACH AND DATA

The data for this analysis were generated in the context of a broader study on science and technology activities across Europe (Noyons et al., 2003a, 2003b). The study referred to three areas of life sciences (genetics, neurosciences, and immunology) as well as nanotechnology. In order to characterise different dimensions of the performance of scientific institutions, publications were taken as a proxy for basic research and scientific orientation, and patent applications as a proxy for applied research and technology orientation. The publication analysis was performed by the Centre for Science and Technology Studies (CWTS) of Leiden University, The Netherlands; Fraunhofer ISI was responsible for the patent analysis.

All of the selected areas can be characterised as knowledge– or science–based. This notion means that the generation of new technology refers to the results of basic research to a large extent. Grupp and Schmoch (1992) and Narin and Noma (1985) demonstrated that patents in science–based areas cite scientific publications more frequently than patents in other areas, thus providing a method to operationalise this feature.

At first, the fields of analysis were defined from a scientific perspective. For the investigation of patents it was necessary to determine areas of potential technological output of scientific disciplines in addition. As to genetics, genetic engineering can be considered as the most important technological application, supplemented by drugs for genetic diseases[1]. The content was described by codes of the International Patent Classification (IPC) and some search terms (Table 32.1).

Table 32.1. Search strategy for patent applications in genetics

IPC codes: C12N015, A01H001, A61K048, A01K067-027, A01K067-033, C12N005-10, C12N005-12, C12N005-14, C12N005-16, C12N005-18, C12N005-20, C12N005-22, C12N005-24, C12N005-26, C12N005-28, C12N001-11, C12N001-13, C12N001-15, C12N001-19, C12N001-21, C12N007-01, A61P037/IC AND ((A61K035 NOT (A61K035-0! OR A61K035-10)), A61K038, A61K039, A61K048, A61K049, A61K051
IPC codes combined with search terms: (C07H021 OR C12Q001-68)/IC AND (GENE OR GENES OR GENETIC? OR GENOME OR GENOMIC?)

\# truncation up to one character (0 or 1);
! truncation of exactly one character (1);
? unlimited truncation (0 or any number).

[1] In the study we exclusively included drugs produced by biotechnical methods.

For the purpose of the present analysis, it is necessary to link the patent applications to scientific institutions. In standard approaches, this is done by a statistical analysis of applicants, i.e., the firms applying. However, the analysis of scientific institutions in a European context is more complex. In most cases the non-university research institutions are registered as applicants and can be examined without major problems, at least at the level of principal organisations. In the case of universities, sometimes:

– the university appears as the applicant;
– the inventor himself, frequently a professor, is the applicant;
– the property rights are transferred to firms which appear as the applicants.

In the latter case the link to the university is only visible through the inventors. As a consequence it is necessary to identify the institutional affiliation of the inventors.

In the European study we determined data sets of patent applications for the priority years 1996 to 2000 and constructed an in-house database, comprising applications directly made to the *European Patent Office* (EPO), called direct EPO applications, and international applications according to the Patent Co-operation Treaty (PCT) with the EPO as the destination office (Euro-PCT applicants)[2]. With regard to genetics, the in-house database covered 21,507 applications of world-wide origin, of which 7,111 had at last one applicant or one inventor from EU or EU–associated countries.

For the determination of the institutional affiliation of the inventors we extracted all inventors from EU and EU–associated countries from the database, all in all 18,930 names. With the support of CWTS these inventors' names were matched to authors' names in the *Science Citation Index* (SCI). For this identification process, several different selection criteria were tested. The following ones proved to achieve the best and most reliable yield:

– surname and initial of the first name;
– identical country of the inventor and author;
– comparable technical and scientific fields;
– identical period of patent application and publication.

The unsystematic use of special letters such as the German ö or the Danish ø proved to be the major problem with regard to the match of surnames; it could be solved in most cases. The search had to be limited to the initial of the first name, as only the initials are recorded in the SCI. Therefore potential sources of error were different first names with identical

[2] The necessary patent records were provided by the EPO.

initials and the unsystematic use of multiple first names. All uncertain cases were excluded from the data set used for matching institutions so as to be on the safe side. The identity of the country of the inventor and the country of the author eliminates potential mismatches to a large extent. In particular, wrong associations occur in the case of countries with the same language (Austria/Germany, Ireland/UK, Belgium/France, Belgium/Netherlands, etc.). An important step is the restriction of the match to scientific fields comparable to the technological fields. Frequent names appear in many disciplines and would imply many mismatches if the searches were executed in the SCI records for all disciplines. As in the study, publications and patents were analysed in parallel, the field of genetics was defined in scientific terms as well, so that the match of inventors could be restricted to the referring authors. Owing to the structure of the SCI database, the search had to be limited to first authors in the case of co-publications of several institutions; only in the case of one institution could all authors be included.

The validation of the matches comprised several steps and is described in Noyons et al. (2003a, p. 29 ff.) in more detail. Supported by various software tools it was primarily based on manual checks of all database entries with regard to inventors and institutions. In addition to the institutional information from the applicant field and the SCI match, further input could be used from the inventor field, as in about 10 per cent of all cases, the inventors record their institutional, and not their private, address. In genetics, about 771 of the inventors' addresses contained such institutional data. In 448 cases, equivalent to 58 per cent, the institutions from the SCI and the inventor field proved to be identical; thus in many cases the correctness of the SCI match could be confirmed[3]. All in all, the number of institutions identified by the SCI match is lower than the real number owing to the exclusion of unclear cases and to the restriction to a specific subject area within the SCI.

3. RESULTS OF THE EUROPEAN ANALYSIS

The analysis of applicants already shows a high share of public research institutions of about 30 per cent. These public institutions may be labelled non-profit institutions in contrast to firms and for-profit research institutes. In this classification privately institutionalised transfer offices working on behalf of universities are considered 'non-profit'. With the introduction of

[3] The share of identical matches is 'only' 58 per cent, because many institutional data in the inventor field refer to firms.

additional institutions by the SCI–based match of authors and inventors, the share of non-profit organisations achieves a level of about 50 per cent in all fields considered (Figure 32.1). Thus the number of non-profit institutions increases substantially by the matching process. The high share even before the match is primarily owed to patent applications from non-university research organisations, also from universities in some countries. By the SCI match we mainly included additional university links which were not visible by using the simple inventor names.

The share of non-profit institutions depends on the method of counting them. In the share of 52 per cent for genetics, the for-profit and not-for-profit organisations receive one count if they appear as an applicant or by an inventor reference. In the specific case of a firm as applicant and an inventor from a university both would be counted once. If we assume that such an invention is exclusively based on university research and that the input from the firm is negligible, an extreme assumption, only the university would be counted. The latter assumption would lead to a share of 63 per cent of non-profit institutions in genetics. In any case, the share of 52 per cent in Figure 32.1 represents a minimum level.

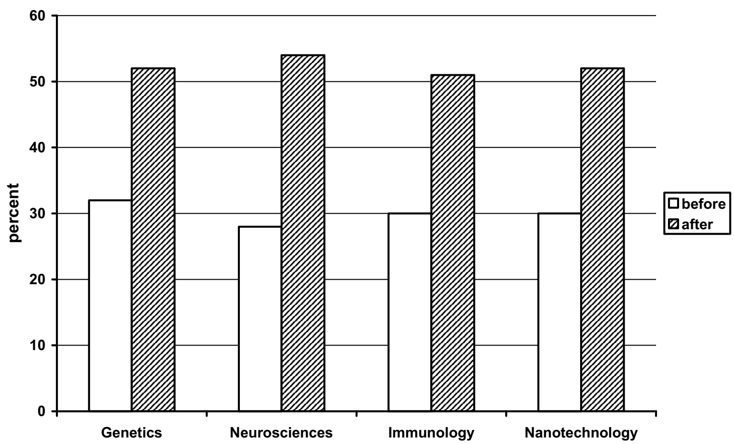


Figure 32.1. Share of non-profit institutions within EPO and PCT patent applications from EU and EU–associated countries in four selected fields before and after the SCI match of inventors and authors, 1996–2000[4]

Source: Noyons et al., 2003a, p. 58–59; Noyons et al., 2003b, p. 32 ff.

The average share of 52 per cent does not apply to all EU countries. Finland, France, Ireland, and Spain are distinctly above this level (Figure

[4] Individual inventors were excluded.

32.2[5], Austria, Denmark, and Sweden have lower shares. The additional input by the SCI match is quite different by country. For instance, the share of non-profit organisations identified by the applicant information is quite low in Austria with 3 per cent, and it achieves 37 per cent with the SCI match. This large difference can be explained by an inventor law, valid in the observation period, which allowed university professors to exploit their intellectual property on their private account, whereas the universities did not appear as applicants. The situation in Germany was comparable. In France the share of non-profit organisations in the applicant information is already quite high, owing to a high weight of non-university research organisations, and early patent activities of universities. In the case of the United Kingdom the considerable difference of the shares before and after the SCI match (30 and 53 per cent) does not meet expectations, as the universities have actively engaged in patenting since the eighties (OECD, 2003, p. 25). The difference is not owed to an inappropriate assignment of external transfer offices of universities to for-profit institutions such as ISIS Innovation Ltd. in Oxford. Obviously the universities often do not lay claim to the inventions of their staff[6], but allow them to exploit their inventions on their private account, or they directly cede the rights to firms

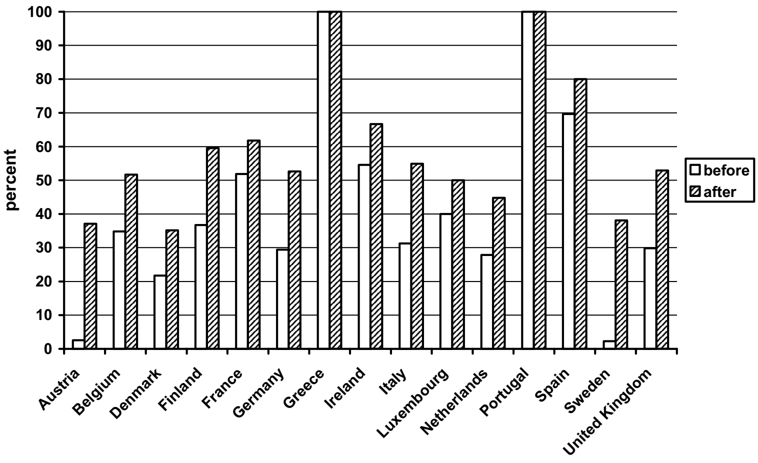


Figure 32.2. Share of non-profit institutions within EPO and PCT patent applications from all EU countries in genetics before and after the SCI match of inventors and authors, 1996–2000[7]
Source: Noyons et al., 2003a, p. 61

[5] The high values of Greece and Portugal are the effect of low absolute numbers.
[6] But they increasingly do so.
[7] Individual inventors were excluded.

So even in the United Kingdom the mere examination of the applicants would be insufficient to reflect the technological orientation of public research institutions. It can be assumed that even with a broader engagement of European universities in patenting, as projected in OECD (2003), the complex match of patent and publication data will be still necessary.

In the project for mapping scientific and technological activities across Europe, patents and publications were analysed as performance indicators in parallel. But we have to ask whether the information on patents is really additional to publications or simply confirms the performance already reflected in publications. For this purpose we exclusively examined a data set of universities in order to achieve a comparable institutional framework. Furthermore we did not include universities with no patents or less than ten patents in the five-year observation period in order to avoid statistical outliers owed to low absolute numbers. The remaining data set of 61 universities (out of a total of 209) was characterised by the relation of publications to patent applications with the assumption that high values reflect a distinct science orientation and low ones a strong technology orientation. In addition, the scientific performance was assessed by the so-called (scientific) impact rate, that is, the specific citation rate of publications normalised by the average citation rate of the field. This simple approach has several methodological shortcomings; the most important ones are:

- As the patent culture at universities is in its beginnings in Europe, their propensity to patent may differ by country and within countries.
- The publication numbers and citation rates between European countries have a positive language bias towards English–speaking countries (Grupp et al., 1999).
- The analysis refers to universities as main organisations. Within universities different research centres with different science and technology orientations may co-exist.

With regard to activity profiles of laboratories, Laredo distinguishes, in the context of human genetics, four types of involvement in:

- “research training
- academic activities...
- industrial activities...
- clinical activities” (Laredo, 1999, p. XI f);

This approach is largely applicable to genetics as well. Based on these features, he found four types of labs[8]:

– labs with "no marked" involvement in specific activities (22 per cent);
– "all embracing" labs (33 per cent);
– "socio-economic only" labs (22 per cent);
– "scientific only" labs (23 per cent) (Laredo, 1999, p. 83 ff.);

In this typology the type "socio-economic" is largely equivalent to a strong technological orientation. According to this typology, only one third of the laboratories prove to be performers in all dimensions. In particular, the orientation to technology or to science seem to exclude each other, at least in many cases (45 per cent). Therefore we can assume that in the university sample for genetics, the science orientation is related to a higher scientific performance.

The graphic representation in Figure 32.3 weakly supports this thesis. However, the picture is quite blurred in the area of lower scientific orientation (indexes between 10 and 25). As a consequence the correlation index is modest with R = 0.29 (significance level of 5 per cent). In Figure 32.3 the countries of the universities are marked at each data point, showing a considerable concentration of British universities (GB) in the area of high technology orientation and high scientific impact in parallel. This may be owing to a high propensity to patent, related to a long tradition in this activity, or an over-assessment by the impact index linked to a language bias. If the British universities are taken out, the correlation improves to R = 0.47 (significance level of 1 per cent) so that the thesis of a certain opposition between a technological and a scientific orientation is confirmed. As a consequence, patent applications and publications reflect different dimensions of the performance of universities.

Although the focus in the European project was on European institutions the US structures were analysed for comparison. For this purpose the US applicant lists in the area of genetic engineering at the EPO were examined for the priority years 1990, 1995 and 2000. All applicants were classified according to the criteria of non-profit and for-profit institutions with private universities and hospitals as non-profit institutions in order to achieve a certain comparability to European structures[9].

[8] He analysed 392 valid responses to a survey.
[9] Again individual inventors were excluded.

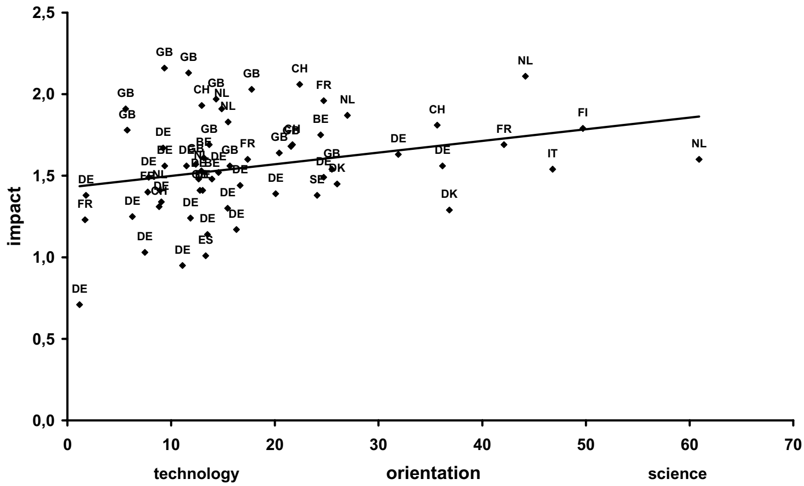


Figure 32.3. Position of European universities according to scientific impact (normalised citation rate) and technological/ scientific orientation (publications/patent applications) in the field of genetics, 1996–2000.
Source: Noyons et al., 2003a; calculation of the author

In the year 2000 the share of non-profit organisations was 25 per cent and thus substantially lower than the European average in the period 1996 to 2000 (52 per cent). The limited comparability of European and US structures may lead to a too low assessment of the US share, but this effect cannot explain the considerable difference; the finding of a clearly lower share of non-profit organisations in the US is still valid. But looking at the priority years 1990 and 1995 this share in the US is much higher with a level of 41 and 42 per cent respectively. Obviously the non-profit organisations took a lead position in the first stage of the technology life cycle of a knowledge–based technology, and in later phases the relative engagement of industry increased. In this perspective the technological development of genetics in Europe lags behind the US development. However, this thesis is a single observation for one specific field; it needs further confirmation by analysing other science–based fields.

4. COMPARATIVE RESULTS FOR GERMANY

The overall share of 52 per cent of non-profit organisations in genetics and similar shares in the other fields analysed are far above expectation. Although the results of the SCI match were checked carefully, doubts remain that methodological errors biased the outcome. Therefore I examined the patent activities of non-profit German organisations at the German Patent and Trademark Office (DPMA), that is, domestic patent applications, with a different approach, in particular without a SCI match. This is feasible, because universities can be identified by the title 'Professor'. In Germany, the title "Professor" exclusively refers to universities; and professors generally indicate their title in official documents such as patent applications. According to a detailed examination in 1996 (Becher et al., 1996), 20 per cent of all German applications with professors as applicants or inventors do not refer to universities, but to other research organisations or enterprises. Therefore in the present analysis the number of university applications was reduced by this factor. According to some expert interviews, the number of university–based patents without professors as inventors or applicants has increased in recent years and achieved a relevant level. Hence the number of university patents is underestimated.

The analysis of non-university institutions was performed by name searches in the applicant field, because these institutions have been active applicants since many years. The investigation focussed on 11 technology fields which proved to be the most science–based ones according to the operationalisation described above. The definition of these fields is documented in Table 32.2; they cover about 40 percent of all applications with German origin.

Table 32.2. Definition of selected science–based technology fields by codes of the International Patent Classification (IPC)

Technology field	*IPC definition*
Biotechnology	C12M, C12N, C12P, C12Q, C12R, C12S
Semiconductors	H01L, B81, G11C
Organic chemistry	C07
Data processing	G06, G10L
Optics	G02, G03B, G03C, G03D, G03F, G03G, G03H, H01S
Telecommunications	G08C, H01P, H01Q, H03, H04B, H04H, H04J, H04K, H04L, H04M, H04N, H04Q
Materials	C01, C03C, C04, C21, C22
Measuring and control	G01, G04F, G04G, G05B, G05D, G05F
Surface technology	B05C, B05D, B82, C23, C25D
Medical technology	A61B, A61F002, A61F009, A61F011, A61H031, A61H039 , A61M, A61N
Polymers	C08B, C08F, C08G, C08H, C08K, C08L

Looking at the shares of the public institutions within all domestic patents, the value in biotechnology is rather high with 39 per cent in the (priority) period 1998 to 2001 (Figure 32.4). However, this share appears to be lower than the 52 per cent in genetics in terms of German EPO and PCT applications according to Figure 32.2. For an appropriate interpretation it has to be taken into account that biotechnology is more broadly defined than genetics; biotechnology also encompasses less science–based sub-fields. In addition, the share of universities is underestimated, as staff without the title professor are not included. All in all, the German data represent a similar order of magnitude and confirm the findings for European data.

The major outcome of the German analysis is the appropriate assessment of the results for life sciences and nanotechnology in relation to other fields. In the perspective of the German data, these areas prove to be extreme cases, because the other science–based fields such as organic chemistry, materials, surface technology, or medical technology reach levels of about 20 per cent. Nevertheless, all science–based fields exhibit a share distinctly above the average level of 7 per cent with regard to all technology fields[10]. The major exceptions are data processing and telecommunications where the public German research infrastructure is weak compared to other countries such as Japan or France. All in all, the thesis of a relevant direct technology contribution of public research institutes still holds.

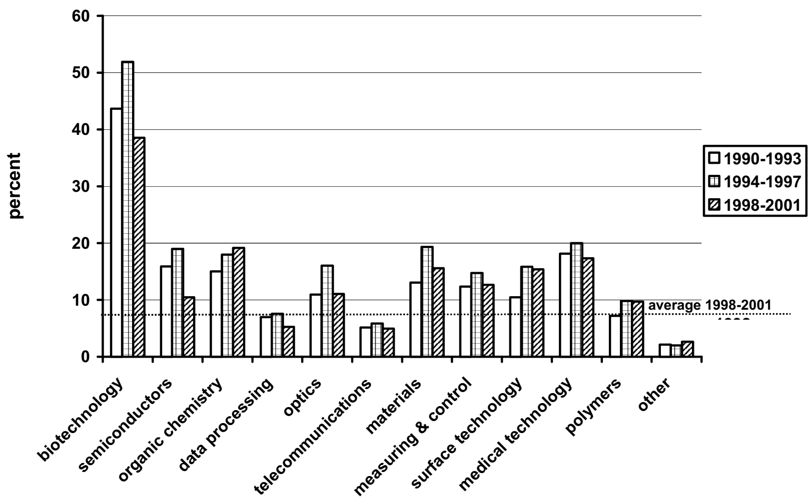


Figure 32.4. Share of German public research institutions within all domestic patent applications with German origin for selected science–based technology fields. Source: PATDPA (STN), own computation

The focus of scientific institutions on science–based fields is reflected in the low share of presently 3 per cent in 'other fields'. 81 per cent of all

[10] The calculation are made for institutions, thus without individual inventors.

applications by public institutions refer to science–based fields, compared to 47 per cent in the case of other actors, in particular, firms.

A further interesting observation is the decline of the public share in biotechnology between the periods 1994 to 1997 and 1998 to 2001, comparable to the decline of US applications in terms of European patents. In contrast, the participation in the 'traditional' field of organic chemistry is steadily increasing. This trend might be an indication of new emerging sub-fields within organic chemistry[11]. These latter statements are, however, reasonable assumptions, but need further substantiation.

5. CONCLUSIONS

If patent applications are taken as a proxy for technological activity it is possible to reflect this specific dimension of the performance of public research institutes. In Europe the identification of patent applications referring to these institutions is complex, because in many cases universities are not registered as applicants. By a match of inventors and authors in the SCI, the identification rate can be substantially increased, leading to a more realistic assessment of the technological activities of non-profit organisations. In the four fields examined in a European study, the share of patent applications of non-profit organisations proves to be quite high. This outcome is confirmed by an analysis of patent applications with German origin in eleven science–based fields in which the share of non-profit organisations is less extreme in most fields, compared to the European results, but still quite high. This finding can be interpreted as a considerable direct contribution to the generation of technology. Therefore the interaction between industrial and public research institutes has to be reconsidered, at least with regard to science–based areas. The various indirect and informal mechanisms of university–industry relations are complemented by a direct contribution of the public institutions to technology generation. In early stages of the technology life cycle of science–based fields, public research institutes play the role of lead actors which is not visible in average measures for all technologies. This direct contribution to technology sheds new light on the function of public research within national systems of innovation.

Only a few research institutes cope with the challenge of being excellent in terms of science as well as technology; rather, the institutes focus their activities on one of these. Generally, a positive relation between scientific

[11] The complete findings of the German study are documented in Schmoch (2004).

performance and science orientation can be found, or a negative one with regard to scientific performance and technology orientation. Patent applications prove to reflect an additional dimension of the performance of public research institutes. As public institutes are increasingly urged to engage in technology transfer, patent indicators are needed to reflect this aspect. However, publication indicators will not become obsolete, because research with a mid- and long-term perspective will remain the major focus of many public research institutes. Technology–oriented research centres will be competitive only if they are able to intensively communicate with science–oriented actors. Therefore more sophisticated indicators are necessary to assess the performance of innovation systems.

On the basis of the above findings, Pavitt's question whether patents reflect the useful research output of universities (Pavitt, 1998) can be answered in a differentiated way: patents reflect an increasingly important dimension of the performance, but should be used in combination with other indicators.

ACKNOWLEDGEMENTS

The research underlying this paper was supported by the European Commission and the German Ministry for Education and Research (BMBF). The author is grateful to the members of the Fraunhofer research team in the European study Sybille Hinze, Rebecca Rangnow, and, in particular, Thomas Heinze for his careful check and verification of the data. Furthermore, I would like to thank Ed Noyons, Renald Buter, and Anthony van Raan of the CWTS team for their fruitful co-operation.

REFERENCES

Becher, G., Gering, T., Lang, O., Schmoch, U. (1996). *Patentwesen an Hochschulen. Eine Studie zum Stellenwert gewerblicher Schutzrechte im Technologietransfer Hochschule–Wirtschaft*. Bonn: BMBF.

Böhme, G., van den Daele, W., Krohn, W. (1978). *The 'Scientification' of Technology*. In W. Krohn, E.T. Layton, P. Weingart (Eds.), The dynamics of science and technology (pp. 219–250). Dordrecht and Boston: D. Reidel Publishing.

Freeman, C. (<1974> 1982). *The economics of industrial innovation*. 2nd edition. London: Frances Pinter Publishers.

Grupp, H., Schmoch, U. (1992a). *Perceptions of scientification as measured by referencing between patents and papers. Dynamics in science–based fields of technology*. In H. Grupp (Ed.), Dynamics of science–based innovation (pp. 73–130). Berlin, Heidelberg and New York: Springer–Verlag.

Grupp, H., Schmoch, U., Hinze, S. (2001). International alignment and scientific regard as macro-indicators for international comparisons of publications. *Scientometrics*, 51 (2), 359–380.

Laredo, P. (1999). *Changing structure, organisation and nature of public sector research. The development of a reproducible method for the characterisation of a large set of research collectives. A test on human genetics research in Europe.* Armines: CSI.

Meyer, M. (2003). Academic patents as an indicator of useful research? A new approach to measure academic inventiveness. *Research Evaluation*, 12, 12–27.

Meyer-Krahmer, F., Schmoch, U. (1998). Science–based technologies: University–industry interactions in four fields. *Research Policy*, 27, Special Issue (Ed. by R. Mayntz), 835–851.

Narin, F., Noma, E. (1985). Is technology becoming science? *Scientometrics*, 7 (3–6), 369–381.

Noyons, E.C.M., Buter, R.K., van Raan, A.F.J., Schmoch, U., Heinze, T., Hinze, S., Rangnow, R. (2003a). *Mapping excellence in science and technology across Europe. Life sciences*. Leiden: CWTS.

Noyons, E.C.M., Buter, R.K., van Raan, A.F.J., Schmoch, U., Heinze, T., Hinze, S., Rangnow, R. (2003b). *Mapping excellence in science and technology across Europe. Nanoscience and nanotechnology*. Leiden: CWTS.

OECD (2003). *Turning science into business. Patenting and licensing at Public Research Organisations*. Paris: OECD.

Pavitt, K. (1998). Do patents reflect the useful research output of universities? *Research Evaluation*, 7 (2), 105–111.

Salter, J., Martin, B. (2001). The economic benefits of publicly funded basic research: a critical review. *Research Policy*, 30, 509–532.

Schmoch, U. (2000). *Wissens- und Technologietransfer aus öffentlichen Einrichtungen im Spiegel von Patent- und Publikationsindikatoren*. In U. Schmoch, G. Licht, M. Reinhard (Eds.), Wissens- und Technologietransfer in Deutschland (pp. 17–37). Stuttgart: Fraunhofer IRB Verlag.

Schmoch, U. (2004). *Der Beitrag öffentlicher Forschungseinrichtungen zur Technikgenese*. Analyse im Rahmen der jährlichen Berichterstattung zur technologischen Leistungsfähigkeit Deutschlands. Accessible through “www.technologische-leistungs-faehigkeit.de”.

Chapter 33

SPECIALISATION AND INTEGRATION

Combining Patents and Publications Data to Map the 'Structure' of Specialised Knowledge

Stefano Brusoni [1,2] and Aldo Geuna[1]

[1]*SPRU, The Freeman Centre, University of Sussex, Brighton, UK;*
[2]*CESPRI and CRORA, Bocconi University, Milano, Italy*
E-mail: a.geuna@sussex.ac.uk

Abstract: This chapter analyses and extends existing studies of how to characterise, trace, and measure knowledge bases of firms, sectors, and countries. The chapter is structured in two main parts. First, we present the concepts of knowledge specialisation and knowledge integration as the relevant dimensions along which knowledge bases can be mapped. The concepts proposed build upon extensive qualitative research which has focused on a variety of *processes* of knowledge generation and use in a range of industrial sectors and organisations. The aim of this chapter is to demonstrate that these, largely qualitative, processes map into key characteristics of the knowledge bases they contributed to generating and shaping; and that these key characteristics can be measured relying on the innovative use of patents, citations and publications data. More specifically, the analysis of the evolution of knowledge specialisation over time provides information about the persistence of knowledge in firms and sectors. It hints at the cumulative, path dependent nature of learning processes. Integration is studied by analysing the evolution of specialisation across different typologies of research. It hints at the complex, non-linear inter-dependence that link the scientific and technological domains. The second part of the chapter will be devoted to the presentation of indicators of breadth and depth that capture the key characteristics of the concepts introduced in the first part.

1. INTRODUCTION

This chapter builds upon the idea that specialisation and integration are two sides of the same coin. On the one side, specialisation is the key process

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 733-758.

through which new bodies of economically relevant knowledge are developed (Pavitt, 1998; Loasby, 1999). On the other hand, specialised competence need to be integrated, or coordinated, in order to deliver more general explanations, or new and better products and processes. Whilst specialisation processes have received quite a lot of attention in terms of developing indicators and methodologies to measure and compare them at various levels of analysis, the development of indicators capable of capturing they key characteristics of those processes which aim at coordinating and integrating specialised knowledge is just beginning to be approached. This chapter is a step in this direction.

The joint analysis of knowledge specialisation and integration processes is relevant for both theory and practice. For example, the integration of dispersed, decentralised knowledge acquisition processes may require further knowledge related investment that allows some firms to act as "loci of coordination." This observation is particularly important for economic analysis because it governs the extent to which knowledge related inputs, at the level of the firm, can be expected to be accurate indicators of the knowledge available to the firm itself. When a firm is able to draw on more extensive networks of knowledge and is able to effectively coordinate these dispersed sources of knowledge generation, traditional approaches to understanding the relationship between inputs and outputs are challenged. Also at country level the ability to coordinate dispersed learning processes increasingly appears to be a key competitive variable for national innovation systems. For example, the study of biopharmaceuticals by Powell et al. (2002) highlights the pivotal and unique role played by the National Institute of Health (NIH) in the US system of research and innovation, and compare it with the more traditional role played by awarding bodies in various European countries, and the EU as a whole.

This chapter is structured as follows. Section 2 identifies the key theoretical developments and concepts that underlie the indicators developed in section 3, at both firm and country level. Section 4 concludes.

2. TOWARDS INDICATORS OF SPECIALISATION AND INTEGRATION

Research in the history of science has stressed the cumulative and social aspects of scientific endeavour. Historians have provided a number of accurate case histories which reveal how the accumulation of results over time influences the rate and direction of the discovery process. For instance, Conant and Nash (1964) describe the process of accumulation of quantitative results in physics that led to Lavoisier's revolution in modern chemistry.

Such a process did not entail the substitution of inaccurate explanations with more accurate ones; rather, it involved the re-conceptualisation of existing findings to deliver a new, more general, explanation. The cumulative development of science has also been intensively studied since the seminal work of Price (1963). On this basis a number of bibliometric indicators have been developed over the years to study the patterns of scientific specialisation at various levels of analysis (micro, meso, and macro). Similar results and indicators have also been derived when analysing the patterns of technological specialisation of large, innovating organisations (Granstrand et al., 1997). Such organisations would appear to follow a rather incremental and cumulative process of accumulation of scientific and technological capabilities. Pavitt (1998) spoke about processes of 'creative accumulation', rather than creative destruction. Indeed, persistence has also emerged as a key feature of specialisation processes. At the micro level it has been accepted for a long time that the learning processes which underpin technical change tend to have become highly routine and history dependent (March and Simon, 1958; Nelson and Winter, 1982). It is easier to learn in the proximity of what one already knows, so to speak. On this basis a huge literature has brought into operational form the concept of knowledge specialisation in terms of the *breadth* of the knowledge base of firms and its evolution, its *persistence* and stability, over time (see, among others, Cantwell, 1989; Grandstrand et al., 1997; Henderson and Cockburn, 1996; Nesta, 2001; Fai, 2003; Brusoni, Criscuolo and Geuna, 2004).

The work of Soete (1987), Pavitt (1989) and Cantwell (1989) provides the building blocks for the analysis of the stability and persistence of technological specialisation patterns at the country level. The line of reasoning is quite simple, yet powerful: if a country is specialised in the 'wrong' (low opportunity) technical or scientific fields one should not expect to be able to refocus one's own specialisation pattern in the short term. Trade and growth indicators will reflect such 'bad' specialisation. Scholars of technical change have therefore devoted much effort to matching technological specialisation indicators and countries' growth indicators (see among others Balassa, 1965; Soete, 1987; Archibugi and Pianta, 1992; Godin, 1994; Dalum et al., 1998; Fagerberg et al., 1999; Meliciani, 2001). This line of enquiry has focused almost entirely on technology (especially patents) and generally do not attempt to provide measurement of the scientific base of the country. The work of Archibugi and Pianta in the early 1990s (Archibugi and Pianta, 1992) is a rare example of the combination of patent studies and bibliometric analysis to examine both scientific and technological specialisation in the EU countries. Figure 33.1 below summarises the discussion above. On the one hand, the breadth of the knowledge bases, which gives an idea of how 'spread' across different fields of knowledge are the competence of a given unit of analysis. On the other side, the persistence of these patterns of specialisation.

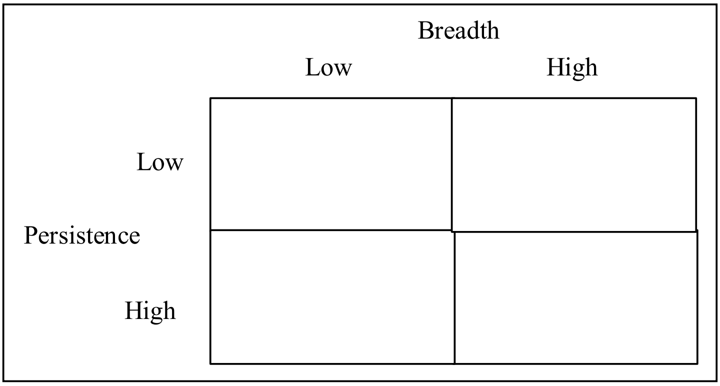


Figure 33.1. Knowledge specialisation: breadth and persistence

The persistence and cumulativeness of specialisation patterns are not the only dimensions relevant to a study of the knowledge bases of firms or countries. Exactly because specialisation is a key mechanism for the growth of knowledge, one should also look at how specialised competence are brought together to deliver more general explanations (the chief concern of scientists), and new and better products and processes (the chief concern of technologists). Indeed, micro level studies of technical change have highlighted how the coordination and integration of different types of competence plays a crucial role in the process of innovation. Integration issues have been studied at length in organisational sciences, strategy, and innovation management literature, beginning at least with the seminal work by Lawrence and Lorsch (1967). For example, Granstrand et al. (1997) studied the distributed capabilities which enable firms to monitor and integrate technologies. Iansiti and Clark (1993) analysed organisational design issues underlying the integration of specialised competence and activities in the mainframe industry. Prencipe (1997) looked at the evolution of the capabilities maintained in house by aero-engine makers to integrate components whose production had been outsourced. Pisano (1997) studied in detail a sample of pharmaceutical development projects in order to conclude that success is related to the ability to carry out, in a co-ordinated and timely manner, a number of heterogeneous, specialised activities which go well beyond the traditional boundaries of the R&D laboratory.

However, whilst the integration of heterogeneous competence seems to be the natural counterpart of the analysis of specialisation processes, the development of indicators that capture some key elements of such a process is only at its very beginning. At the micro level we were able to identify only a few works which aimed at developing measures of such integrating processes. Building upon the seminal work by Rumelt (1974) and Teece et al. (1984) the concept of coherence has received quite a lot of attention in the literature, both in theoretical and empirical terms. For example, Nesta (2001) measures the integration of the knowledge base of firms in terms of their technological coherence. Coherence is a measure of the relatedness of firms' technological capabilities. A number of indicators has been developed based upon various definitions of 'distance' to analyse how 'coherent' the specialisation patterns of innovating organisations are. The work of Nesta (2001) has initiated a wide exploration programme of this concept, its measures, and its impact over firms' innovative and business performance (Nesta and Saviotti, 2003; Nesta and Dibiagio, 2004). In a way, one might argue that coherence facilitates the integration and coordination of specialised competence by reducing the cognitive and behavioural distance that separate those communities of scientists and engineers who develop specialised fields.

The second dimension of integrating activities we found in the literature is related to the complexity and criticality that characterise specific fields, or components. This is what Wang and von Tunzelmann (2000) refer to when they talk about the increasing analytical complexity, or technical difficulty, of specific fields of research. It is in this respect that they talk about 'depth'. Very few indicators of depth exist. Below we focus on those few examples we know. The reasoning underlying the development of depth as an indicator of integration is simple. Complex fields of research are those in which a clear cut way of decomposing a problem to sub-problems is not known. Or, if known, sub-problems and tasks are linked by unpredictable interdependence (Perrow, 1967; March and Simon, 1958). In order to manage such complex interactions it is necessary to maintain capabilities which span all the sub-problems, tasks, and activities concerned. In practical terms it means that should a firm be involved in R&D activities focused on a new, complex field of research (say, neurosciences or fluid dynamics) such a firm would be required to be engaged in basic, applied, and engineering oriented research in order to be able to evaluate the soundness of any candidate solution. Hence, the indicator we propose below looks at patterns of specialisation (within one specific field of research) across typologies of research. For example, Prencipe (2000) studied depth in the case of the evolution of the aero-engine control system. His study focused on two dimensions of the knowledge bases of aero engine makers: the stages of the development process performed by engine makers, and the different types of knowledge related to either the combination of control system components, or specific components (p. 898). His study showed the importance of considering this additional dimension (depth) alongside breadth. Henderson (1994) analyses the role played by capabilities of integrating in pharmaceuticals R&D. Brusoni, Criscuolo, and Geuna (2004) attempt to quantify integration in term of the depth of firms' knowledge bases using publicly available patent data (see also below).

Very little attention has been devoted instead to the measurement of integration at the country level. To our knowledge the only work which attempts to do so has been produced by the authors of this paper (see Brusoni and Geuna; 2003). The knowledge integration of a country has been determined in terms of the depth of its knowledge base as measured by the specialisation across research types (i.e. basic, applied and engineering oriented research). Figure 33.2 summarises this discussion. Integration activities can be captured by using indicators of coherence and depth. These two indicators capture different mechanisms of integration. Coherence looks at the 'distance' between technological subfields. Depth looks at the involvement in different types of research for a given field.

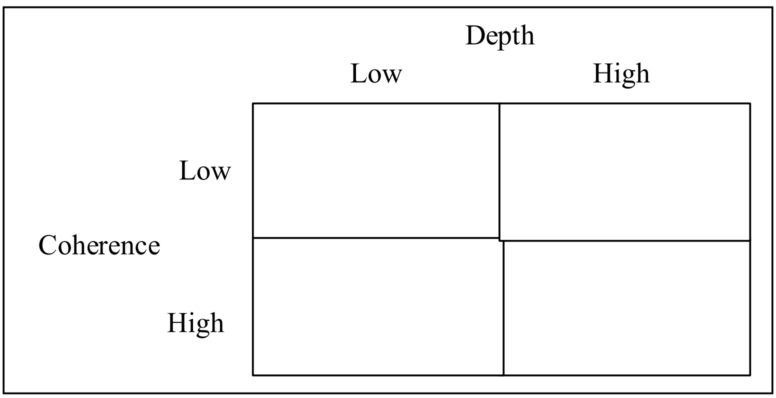


Figure 33.2. Knowledge integration: depth and coherence.

3. MEASURES OF KNOWLEDGE SPECIALISATION AND INTEGRATION

Different ways of measuring knowledge specialisation and integration have been developed in recent years both at the micro and macro level of analysis. Drawing upon Brusoni, Criscuolo, and Geuna (2004) and Brusoni and Geuna (2003), this section focuses on two examples, one at the company level and one at the country level, of how to operationalise the concepts of knowledge specialisation and integration. Also, we will discuss how these measures can contribute to the ongoing debate on the knowledge production and distribution in the process of economic development. Both examples focus on the pharmaceutical industry, a fast growing, rapidly changing sector that in the past three decades has greatly increased both its contribution to economic growth and its visibility in the public policy arena.

3.1 The Company Level Analysis

The process of innovation in pharmaceuticals companies has been undergoing significant changes owing to the developments in biology and genetics and to the emergence of biotechnologies. These changes are connected with modifications in the knowledge bases of large pharmaceuticals companies. Is it possible to define ways of measuring the structure and organisation of the knowledge bases of these firms? To do so we operationalise the concepts of knowledge specialisation and integration with measures of the breadth (and its evolution) and depth of the knowledge bases of the 30 largest pharmaceutical corporate groups in the world. The patent portfolio of these companies is used as the source of information for calculating a set of proxies of the breadth and depth of their knowledge bases.

Patent documents contain citations to other patents and references to a variety of other publications such as papers, abstracts, conference proceedings, books, etc. (non-patent references), to fulfil the legal requirement of supplying a complete description of the state of the art. Citations limit the scope of the inventor’s claim to novelty and, in principle, they represent a link to previous innovations or existing knowledge with reference to the innovation. Non-patent citations can be used as a proxy for some characteristics (breadth and depth) of the knowledge bases upon which firms build. They hint at the bodies of scientific and technological knowledge on which patents (and therefore firms) rely, but, for the current purposes, are not meant to imply any direct links between the inventors and researchers active in those fields. They are not used to imply any specific network structure, or localisation effects. They point to bodies of knowledge which the inventors (or the examiners) thought relevant to the invention that is the subject of the patent application[1].

We use an original database that includes the 33,127 patents filed with the EPO by the 30 largest corporate groups in the pharmaceuticals industry during the period 1990 to 1997 (consolidated in 1997). From the 33,127 patents 41,931 non-patent references were extracted. Out of these we identified 25,996 citations to scientific articles included in the expanded ISI *Science Citation Index* (SCI). We decided to focus only on the scientific publications included in the SCI database for three reasons. First, the SCI database includes publications in peer reviewed journals with international recognition; second, journals in the SCI database are classified in 132 scientific fields; and third, journals in the SCI database can be linked via the Computer Horizons Inc. (CHI) journal classification to four major categories of research - applied technology, engineering technological science, applied research and basic research. In the context of biomedicine the four types of research levels are: clinical observation (Level 1), a mix of clinical observation and clinical investigation (Level 2), clinical investigation (Level 3), and basic research (Level 4) (Narin and Rozek, 1988). As Narin and Rozek state, Level 1 is typified by the *Journal of the American Medical Association*, Level 2 by *The New England Journal of Medicine*, Level 3 by the *Journal of Clinical Investigation*, and Level 4 by the *Journal of Biological Chemistry*.

Clearly the use of the SCI scientific fields[2] classification and CHI classification have some limitations, which are thoroughly discussed in the literature. Notwithstanding these limitations, the SCI scientific fields classification can be used to measure which scientific fields were of relevance for the patents of the groups and therefore can be used to calculate a set of proxies of the breadth of the knowledge base upon which patents (and therefore the company) rely. Whilst the CHI classification does allow the non-patent citations to SCI scientific journals to be classified in research types. On this basis we can measure which research typology was of relevance for the patents of the groups and use this to develop a proxy for the depth of the knowledge base of the company in terms of its reliance (or not) upon research types which are covering the spectrum from basic research to clinical observation (versus being focused only on one type be it basic research or development).

1 It is, of course, true that the choice of citations to non-patents literature may also be affected by strategic considerations.

2 All SCI fields are reported in Appendix 1 below.

3.1.1 The breadth of the knowledge base

The breadth of the knowledge base of a company can be calculated in absolute levels (how many citations in each scientific field) or in terms of its relative specialisation in certain scientific fields. Both approaches provide some information on the breadth of the knowledge base of a company in terms of the presence in certain scientific fields and the company relative specialisation in others.

Table 33.1 presents a summary of the individual specialisation profiles of 10 selected corporate groups; it includes the top six scientific fields by number of citations, the number of scientific fields with at least 10 citations (over the whole period) and the scientific fields that have a symmetric Relative Specialisation Index (RSI) > 0.3 defined as *core scientific specialisation* fields.

The symmetric RSI, derived from the Revealed Technological Advantage index (Balassa, 1965; Soete, 1987), is obtained standardising the activity index (AI), which is defined as the share of citations in a given scientific field in the citation portofolio of a given corporate group relative to the share of citations in a given scientific field, for all corporate groups, in the overall sample of citations[3]. The symmetric RSI index indicates whether a firm has a higher-than-average activity in a scientific field (RSI > 0) or a lower-than-average activity (RSI < 0).

$$\mathrm{AI} = \frac{\dfrac{p_{ij}}{\sum_i p_{ij}}}{\dfrac{\sum_j p_{ij}}{\sum_i \sum_j p_{ij}}} \qquad\qquad RSI = \frac{AI - 1}{AI + 1}$$

where p = number of cited publications, $i = 1 \ldots n$ = number of scientific fields = 132 and $j = 1 \ldots m$ = number of corporate groups = 30.

Finally, we define *distinctive scientific specialisation* as those fields among the top six for which the corporate group has a symmetric RSI > 0.3. On average, the groups have at least 10 citations across 18 scientific fields.

[3] In calculating the RSI index we used a threshold level of 10 citations in the period under analysis. However the two totals for the group and for the overall sample include citations in all scientific fields regardless of the specified threshold level.

For example, Hoechst AG (43), Novartis (39) and Bayer AG (35) are active in more than 35 fields, while Teijin Ltd (3) and Yamanouchi Pharmaceuticals Co. Ltd (4) have less than six fields with at least 10 citations over the whole period.

Table 33.1. Top six scientific fields and core scientific fields

Corporate Group	*Top 6 scientific fields*	*No. Fields*	*Core scientific specialisation, RSI>0.3*
Abbott Laboratories	CQ DX DY NI QA RO	20	<u>QA</u>, CO, EA, JY, NN, PW, ZE
Baxter International	CQ DQ MA NI QA ZD	6	<u>DQ, MA, QA, ZD</u>
Bayer	CQ DX DY EE TU UY	35	<u>EE, UY</u>, AM, DW, EA, EC, EI, II, IQ, IY, OA, PM, PW, SY, UB, UE, VY, YE
Hoechst	CQ DX DY EE RO TU	43	CU, EC, EI, II, IY, PK, PM, PW, UB, UK, UY, WE, YA, YP, ZD
Monsanto	CQ DE DR DY EE RO	12	<u>DE, EE</u>, AM, DW, UY
Novartis International	CQ DE DY EE RO TU	39	<u>DE</u>, AM, DW, EA, EC, EI, II, IY, JY, KM, MU PM, UY, VE, YA, YP, ZD
Teijin	CQ MA RO	3	<u>CQ, MA, RO</u>
Yamanouchi Pharmaceutical Co.	CQ DX DY TU	4	<u>DX, DY, TU</u>
Zeneca	CQ DE DX DY EE TU	15	<u>DE</u>, AM, CU

No Fields: Number of scientific fields with at least 10 citations. <u>Underlined</u> are the fields of positive specialisation that are also within the top six fields, we consider these the *fields of distinctive specialisation*. All SCI fields are reported in the Appendix at the end of this chapter. For the table with 30 corporate groups see Brusoni, Criscuolo and Geuna (2004).

Although the top six fields across groups tend to be quite similar, some important differences appear both in terms of field concentration and field specialisation. A few corporate groups stand out in terms of their focus on certain scientific fields. For example, the top six scientific fields for Baxter International are biochemistry and molecular biology, cardiac and cardiovascular systems, haematology, immunology, medicine (research and experimental) and peripheral vascular diseases. Also, Baxter International has very strongly defined *distinctive scientific specialisations*, with only four fields with symmetric RSI > 0.3 all of which are in its top six fields. This specialisation profile seems to point to a specific market focus for Baxter's innovative activities. Although most groups do not show such a focused specialisation profile, nevertheless a few specific scientific fields, relatively important in their citation portfolios, can be detected. For example, Monsanto has a *distinctive scientific specialisation* in plant sciences and organic chemistry, these being among the top six fields with a symmetric RSI of respectively 0.77 and 0.36. Similarly, Zeneca, although with a large number of active scientific fields, has a clear specialisation profile with an important *distinctive scientific specialisation* in plant science (top six and RSI = 0.8) and *core scientific specialisation* in agronomy and biology. At the other extreme Bayer AG, Hoechst AG and Novartis, with large citation portfolios, show a more diversified pattern of scientific specialisation making it quite difficult to characterise in a clear cut manner their scientific profile in terms of specific competences.

From the results of this preliminary analysis it is clear that the 30 groups rely, at least in part, on different knowledge bases (for example, both Bayer AG and Novartis have very broad specialisation profiles yet they exhibit different *distinctive specialisations* in, respectively, polymer sciences and plant sciences). However, analysing non-patents citations across the whole period provides only limited information on the breadth of the knowledge bases of different corporate groups. In order to overcome this limitation, we focus on the evolution of the breadth of the knowledge bases over time.

3.1.2 The evolution of the breadth of the knowledge bases

To study the persistence (or lack thereof) of the knowledge bases of the groups considered, we examined their citation portfolios in two sub-periods: 1990-92 and 1995-97. We compared these two periods with the aim of examining trends in the evolution of the breadth of the knowledge bases. Owing to limitations of space we report only a few of the possible indicators of the evolution of breadth.

First, we looked at which fields each group was active in during both periods (persistence), and which fields each group cited at the end of the period only (entry), or at the beginning only (exit). Also, measures of similarity and concentration by sub-period were calculated for each group citation profile. Second, we calculated the symmetric RSI for the two periods and examined the entry and exit in fields of *core scientific specialisation*.

For each of the groups we calculated the changes in the breadth of the knowledge bases in terms of entry, exit and persistence in the citations to scientific fields. As an example, Table 33.2 shows in the case of groups the number of scientific fields in which each group was active in the second sub-period but not the first (entry: number of new fields in 1995-97)[4]; the

[4] 'Active' indicates that a group has cited at least 10 times articles classified in a particular scientific field in a given sub-period. It is worth pointing out that this means that the

number of fields in which the group no longer has citations (exit: number of fields exiting); and finally the number of fields in which the group cites in both periods (number of persistent fields in the two sub-periods).

Table 33.2. Entry, exit and persistence in scientific fields

Corporate Group	*No. active fields throughout period*	*No. new fields in 1995-97*	*No. fields exiting*	*No. persistent fields in the two sub-periods*	*Similarity measure*
Abbott Laboratories	20	5	1	8	0.81
Bayer	35	10	1	7	0.99
Rhone-Poulenc	27	12	0	9	0.37
Teijin	3	0	1	0	0

No Active Fields: Number of scientific fields with at least 10 citations.
For the table with 30 corporate groups see Brusoni, Criscuolo and Geuna (2004).

To examine the changes in the breadth of the knowledge base of a group in terms of the number of scientific fields in which it is active, we used a measure of similarity originally derived by Jaffe (1986). In our case the indicator (S_k) provides a measure of the scientific distance between the breadth of the specialisation profile of each group in the two sub-periods. The distance (variation in breadth) in scientific specialisation across time can be approximated by an uncentred correlation coefficient of the vectors (f_i and f_j) of citation share in each scientific field for each group in the two sub-periods[5].

$$S_k = \frac{f_i f_j^{'}}{\sqrt{(f_i f_i^{'})(f_j f_j^{'})}}$$

where k = 1…30; corporate groups and i = period 1990-1992; j = period 1995-1997.

---

number of active fields throughout the period is in general higher than the sum of the number of persistent, exiting and new fields. There might be, in fact, fields with more than 10 citations between 1990~97 but those citations are uniformly spread across the whole period and they never reach the threshold level in any one particular sub-period.

5 The uncentred correlation coefficient (or cosine measure) was preferred to the standard correlation coefficient because whilst the first measures the cosine of the angle between the two observation vectors measured from the mean, the second measures this from the origin (zero citation share in a scientific field) of these vectors.

This similarity measure is bounded between 0 and 1, and the greater the degree of similarity between the breadth of the groups in the two periods the closer it is to unity.

In terms of similarity, most groups show a high level of similarity (an average of 0.75). 21 groups have a similarity index above 0.5; of these 14 have an index above 0.8 and seven have an index above 0.9. However, four groups do not have any overlap amongst the fields cited at the beginning and at the end of the period (these are the same groups with a small number of citations, which have less than six fields with at least 10 citations), and five groups have a similarity index lower than or equal to 0.5. The table shows a weak trend toward increasing the range of fields in groups' citation portfolios. Taking the arithmetic average of the differences between entry and exit, the result is an average 'net entry' for two fields (2.23). Adjusting this average by the number of fields in which groups are active throughout the period (calculating the ratio between net entry and total number of cited fields over the entire period, then taking the arithmetic average), one gets a positive increase in breadth (0.12).

Given these results we proceeded to the examination of the changes in the specialisation profiles of the groups. To do this we calculated the symmetric RSI for all groups and for all fields for the initial period (1990-92) and the final one (1995~97). We focus on entry and exit into fields of core scientific specialisation. Table 33.3 builds upon Table 33.2 focusing on the fields which groups have ceased to cite or have begun citing, whose symmetric RSI is above 0.3. In other words, we look at changes in the groups' fields of core specialisation.

Table 33.3. Entry and exit in *core specialisation* fields

Corporate Group	*Fields exiting*	*New fields*
Abbott Laboratories	PW	CO, DB, KM, QU ZE
Bayer		AM, EA, EI, PW, SY, VY
Bristol-Myers Squibb	DB, DQ, PY, QU, RU, VE, WE, ZD	DM, FQ
Glaxo Wellcome	FQ, KM, NN	DM
Hoechst	II, UB, UK, ZD, ZE	CU, DW, EA, EI
Novartis International	IA, PY, RT, RU, VE	CO, DQ, DW, EA, EI, II, MA, YP, ZD
Sanofi Synthelabo	DB, DE, JY, KM, QU	QA, RU

Only scientific fields with symmetric RSI > 0.3. All SCI fields are reported in the Appendix at the end of this chapter. For the table with 30 corporate groups see Brusoni, Criscuolo and Geuna (2004).

Again, cross-group heterogeneity emerges quite visibly. However, some trends are also observable. Of the 30 groups considered, 14 show no exit from the fields of core specialisation (those with a symmetric RSI above 0.3). Seven groups show the loss of only one field of core specialisation. Five groups (Bristol Myers Squibb, Glaxo Wellcome, Hoechst AG, Novartis and Sanofi-Synthelabo) exhibit a fairly big change in core specialisation with four or more fields exiting the core. Conversely, on the entry side only five groups show no new entry in the areas of core specialisation, eight groups have at least one new entry, and seven groups exhibit at least four new fields with a symmetric RSI above 0.3. Bristol Myers Squibb, Hoechst AG and Novartis appear to be the most active groups in terms of both entry and exit. In some cases we can distinguish a few clear patterns.

For example, it seems clear that Abbott Laboratories is moving into the bio-pharma area: new fields of core specialisation include biochemistry research methods, biotechnology, genetics and heredity, microbiology and virology. In other cases the patterns are less clear. For example, Bayer seems to be moving into traditional fields such as analytical chemistry, physical chemistry, and agronomy, but also into optics and radiology. Overall then, pharmaceutical groups appear to be more 'active' on the entry than on the exit side. This is consistent with the argument that the emergence of new fields of useful research leads to an increase in the breadth of the knowledge base.

Overall, the analysis of the knowledge breadth of the largest pharmaceutical groups reveals a high level of heterogeneity at the level of the group in the field of active involvement and in the fields of specialisation when considering either the entire period or the changes from starting to the end. Particularly interesting are the results of the analysis of change in the knowledge bases. Although groups have a high level of similarity in their portfolio of citations at the start and end of the 1990s, we found some evidence of a weak but positive increase in their breadth, a significant entry in new fields of core specialisation, and a low level of persistent specialisation for the stable fields. The analysis allows a few clear patterns at the level of the individual group to be highlighted. For example, whilst Abbott Laboratories is specialising in bio-pharma, Bayer seems to be maintaining, and actually reinforcing, its specialisation in traditional fields.

3.1.3 The depth of the knowledge base

In order to analyse how integrated firms are across research categories (depth of the knowledge base) we calculated an indicator of depth based on the symmetric RSI. The indicator presented in this chapter uses a symmetric RSI calculated for each pair of scientific field and typology of research. Also owing to data constraints (a threshold of at least five citations in each pair), the four CHI categories have been reclassified in two broad types of research to analyse if a group is not only active in the broadly defined applied part of the research (clinical observation and investigation), but it also has competences in the area of basic research. Depth (the indicator of knowledge integration) is measured by the ratio between the numbers of pairs of scientific fields and typology of research that show a positive specialisation in both CHI levels and the number of scientific fields of positive specialisation. So, for example, let us consider the case of 'neurosciences'. In order to be able to say that a company is 'integrated' in neurosciences, we calculate the symmetric RSI for that company on the basis of the citations to neurosciences journals that focus on *basic research*. Then we calculate the symmetric RSI in neurosciences for the same company on the basis of the citations to neurosciences journals that focus on *clinical observation and investigation* if the company considered is relatively specialized (RSI > 0) in both types of research within neurosciences, then we can say that this company is 'integrated' in neurosciences, and we add 1 to the numerator. The numerator of the indicator then shows in how many disciplines a firm is relatively specialised in both categories of research. The denominator controls for the size of the groups' citation portfolios. Thus depth is derived using the formula below and it varies between 0 and 1. It is 0 when the group considered does not exhibit any overlap between the two types of research. It is 1 when the group considered is fully integrated across all types of research in all the fields in which it exhibits positive specialisation. It represents the share of integrated scientific fields relative to all fields with positive RSI. Those disciplines that were classified in only one of the two CHI categories are excluded from the calculation of depth.

$$x = \begin{cases} 1 & \text{if } \mathrm{RSI}_{\mathrm{i,k}} > 0 \ \ \forall k = 1,2 \\ 0 & \text{otherwise} \end{cases}$$

$$y = \begin{cases} 1 & \text{if } \mathrm{RSI}_{\mathrm{i}} > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\mathrm{DEPTH}_{\mathrm{j}} = \frac{\sum_i x_i}{\sum_i y_i}$$

where $i = 1..132$ number of fields $j = 1..30$ number of corporate groups $k = 1, 2$ number of research types.

Table 33.4 presents the depth indicator and the scientific fields in which the corporate groups are integrated. On average the groups have a low level of depth (0.092) representing about 10% of scientific fields with positive RSI being integrated. Seven groups do not show any integration. Only three corporate groups have DEPTH ≥ 0.2, i.e., Eli Lilly and Co., Rhone-Poulenc SA and SmithKline Beecham plc. Abbott Laboratories and American Home Products also have high values. Eli Lilly and Co. (0.25) has a positive specialisation across research types in five fields: biotechnology and applied microbiology, medicinal chemistry, research and experimental medicine, neurosciences, pharmacology and pharmacy. Rhone-Poulenc SA is integrated in biochemistry and molecular biology, biotechnology and applied microbiology, genetics and heredity, immunology, microbiology and neurosciences. SmithKline Beecham plc is integrated in biochemistry and molecular biology, microbiology, neurosciences, pharmacology and pharmacy. Abbott Laboratories exhibits a fairly clear profile in bio-pharmaceuticals. The integrated fields are biochemical research methods, biotechnology and applied microbiology, microbiology and virology. The extent of overlap with Table 33.3 is remarkable: the integrated fields are all fields that became of *core specialisation* for this group only in 1995-97. Monsanto is integrated in plant sciences only, which reflects the strategic focus of this group over the period considered. Moreover, Hoechst AG and Bayer are integrated, respectively, in only material sciences and multidisciplinary chemistry, reflecting that these groups (in the period considered) were still behaving like traditional 'chemical' companies. Overall, the fields related to the new bio-pharma trajectory seem quite likely to be developed in an integrated manner.

Table 33.4. Depth indicator and code of fields integrated

Corporate Groups	*DEPTH*	*Fields*
Abbott Laboratories	0.18	CO, DB, QU, ZE
American Home Products	0.16	CQ, IA, NI, QU
Baxter International Inc.	-	
Bayer	0.03	DY
Eli Lilly and Co.	0.25	DB, DX, QA, RU, TU
Hoechst	0.03	PM
Monsanto	0.11	DE
Rhone-Poulenc	0.27	CQ, DB, KM, NI, QU, RU
SmithKline Beecham	0.23	CQ, QU, RU, TU

All SCI fields are reported in the Appendix at the end of this chapter. For the table with 30 corporate groups see Brusoni, Criscuolo and Geuna (2004).

In other words, the data in Table 33.4 point to the fact that those groups that have entered this research trajectory have done so by quickly developing strong competence across research typologies in the relevant fields. The data capture the delay of the German groups, and hints at the possibility that such delay is characteristic only of the German groups, rather than all the EU countries. Finally, and from a methodological perspective, we should point out that the depth indicator employed in this section makes it possible to discriminate between groups in a more straightforward way than the breadth indicators calculated above. Specifically, the depth indicator is not biased by the size of the patenting activity of a group, and it provides a better proxy for the strategic research orientation of the group; for example, Bayer AG, Hoechst AG and Novartis have a very high number of patents and citations, but they have very few integrated fields, with the German groups focused on traditional chemicals and the Swiss group integrated in the bio-tech, microbiology area.

3.2 The Country Level Analysis

The country approach aims to identify and operationalise, at sectoral level, the relevant dimensions that make the comparison of the knowledge bases of different countries a meaningful exercise. Particular attention is devoted not only to examining whether each country's specialisation is stable over time (knowledge persistence), but also to whether specialisation by field is similar across different typologies of research (knowledge integration).

The operationalisation of these two dimensions is based upon the design of a comprehensive data set of peer reviewed papers which was obtained by combining the standard SCI classification by science field with the CHI classification of typologies of research. The result is an original data set encompassing some 630,000 papers in 11 different sub-fields of chemistry and pharmacology published between 1989 and 1996 by researchers in the four largest European countries (the UK, Germany, France, and Italy), the EU as a whole, the US and Japan (CHEMPUBS database). This data set is analysed in combination with the PACE (Policies, Appropriation, and Competitiveness in Europe) survey (Arundel et al., 1995). The results of the PACE questionnaire pinpoint the pharmaceutical industry as being a highly internationalised industry. PACE data show that not only do EU R&D managers in the pharmaceutical sector value the results of public research, but also that they rely upon international research much more than those in the chemical sector and in other manufacturing industries. Also, PACE stresses that the pharmaceutical industry relies more on North American research than on EU research. The questions which demand explanation are why do EU pharmaceutical firms rely to such a great extent on North American research? What makes it attractive to EU firms? In attempting to answer these questions, we discuss some evidence related to the existence of a 'European Paradox', a high quality science base, measured through the rate of publication, which does not correspond to a strong technological and economic performance, measured through patenting activity (European Commission, 1997) in the case of traditional pharmaceuticals.

We compare sectoral knowledge bases across countries by developing a grid designed along two of the dimensions identified above in Figures 33.1 and 33.2: depth and persistence. Following the same methodology used in the company level analysis we can operationalise the persistence in knowledge specialisation with an analysis of the evolution of the breadth of the chemistry and pharmacology knowledge bases. In particular, we regressed the symmetric RSI in 1996 on the 1989 value, country by country. If the β coefficient is equal to 1 then the country specialisation pattern has remained unchanged over the period. If $\beta > 1$ then the country is increasing its positive specialisation in fields where it was already specialised. If $0 < \beta < 1$ the country has decreased its non-specialisation in those fields where it was negatively specialised at the beginning of the period (or decreased its positive specialisation where it was positively specialised). In all cases variations in specialisation occur in a cumulative way as $\beta > 0$. In the case that β is not significantly different from zero, the hypothesis that changes in specialisation are either not cumulative or are random cannot be excluded. If β is negative we are witnessing a process of reversion in the specialisation. The case where $\beta >1$ is often referred to as β-specialisation (Dalum et al., 1998). As all coefficients except those for Germany (whose β equals 1) are $0 < \beta < 1$ (β-de-specialisation), we set 0.5 as the threshold dividing high or low persistence in knowledge specialisation.

Similarly, we used the indicator of depth presented above to try to capture the level of knowledge integration of the scientific knowledge base of countries in the fields of chemistry and pharmacology. Differently from the company example, given the large size of the CHEMPUBS database we were able to use a three level classification of types of research, *clinical observation*, *clinical investigation*, and *basic research*. If the country considered has a relative specialisation (RSI > 0) in all the three typologies research in one of the 11 scientific fields, then we can say that this country is 'integrated' in that field (numerator). The denominator controls for the size of the country field specialisation. The threshold between high and low depth is given by the arithmetic average of the indicator (0.22). Figure 33.3 reports the result of such a combination.

It is fairly apparent that the US and Germany combine high levels of both knowledge integration and knowledge persistence. France, despite a high level of depth, exhibits low persistence over time. Neither Japan nor the UK shows any integration, but the breadth of the UK is more stable. Italy and the EU are somewhere in between. The EU as a whole is characterised by both average integration and low persistence (this latter coefficient was not significant in the regression). Italy appears to be relatively integrated, but exhibits low persistence (Italy's coefficient for persistence is not significant).

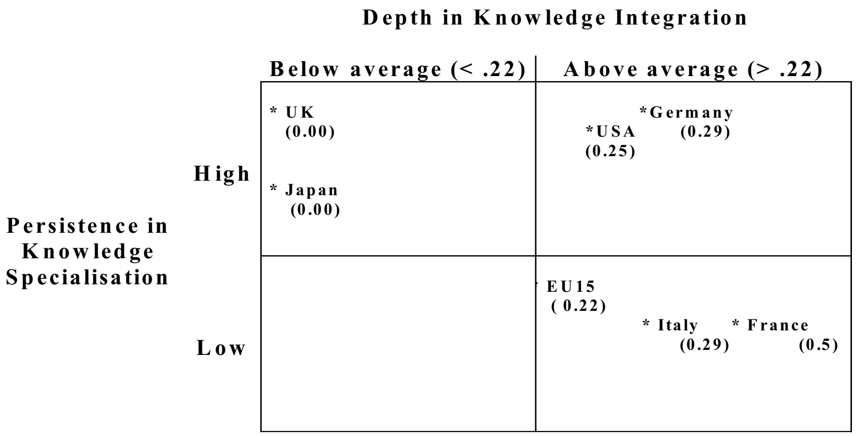


Figure 33.3. Matrix of knowledge depth and persistence
Source: Brusoni and Geuna (2003)

Table 33.5. Fields of positive specialization by type of research

	Applied Technology & Engineering	*Applied Research*	*Basic Research*
EU	C10 C3 [C8]	C4 C7 [C10 C8 C2 C1]	C2 C5[C4 C8 C6 C7]
F	C8 C4 [C6 C10]	C7 C4[C8 C6 C9]	C7 C4 [C8]
D	C3 C4 [C11]	C1 C5[C4 C9 C8]	C5 C4 [C8]
I	C10 C6 [C8]	C7 C6 [C10]	C2 C6 [C7 C5 C4]
UK	C3 C10	C4 C7[C10 C3 C6]	C6 C4[C5 C2]
US	C1 C2 [C9 C6 C10]	C6 C10 [C3 C1]	C6 C10 [C2 C8 C3 C7]
J	C4 C11	C10 C9 [C1]	C1 C10 [C7]

Source: CHEMPUB database — Brusoni and Geuna (2003); elaboration of SCI and CHI data. Top two positive specialisation fields out of brackets.
C1 General Chemistry, C2 Analytical Chemistry, C3 Applied Chemistry, C4 crystallography, C5 Inorganic and Nuclear Chemistry, C6 Medical Chemistry, C7 Organic Chemistry, C8 Physical Chemistry, C9 Polymer Science, C10 Pharmacology & Pharmacy, C11 Chemical Engineering.

The results of this taxonomy should be interpreted together with the analysis of the breadth of specialisation of the countries (see Table 33.5). Despite the high persistence and depth exhibited by both the US and Germany, their specialisation profiles appear to be very different. In particular, Germany's specialisation revolves around traditional chemistry fields, such as crystallography (C4) and inorganic chemistry (C5). The US is specialised in those fields more directly related to pharmaceuticals: medical chemistry (C6) and pharmacology and pharmacy (C10). The other EU countries studied also are more specialised in 'chemistry for chemicals', rather than pharmaceuticals. Furthermore, it is evident from the regressions we ran by type of research that the EU countries' specialisation in medical chemistry and pharmacology decreases as we move away from the development type of research towards applied and then basic research.

Such results are consistent with other studies of specialisation that rely on traditional methodologies.

So for Germany specialisation in 'traditional' chemistry (inorganic and organic) is confirmed by Sternberg (2000, p. 98) who also highlights the German disadvantage in medical sciences. OST (1998) confirms both the integration of the German pattern of specialisation and its focus on chemistry. Furthermore, the UK seems to be more specialised in medical research than France or Germany.

OST (1998) also confirms the strong EU specialisation in chemistry and its relative disadvantage (in terms of publications) in biology (basic research).

These different specialisation profiles hint at a possible explanation for the PACE questionnaire results. The PACE survey revealed that public research carried out in North America was valued and used extensively (even more than public research carried out in other European countries) by the largest EU R&D firms in the pharmaceutical sector. The PACE questionnaire does not allow speculation about why this happens, however. We argue that the reliance of EU firms on the North American knowledge base is consistent with the US' exhibition of a persistent, as well as an integrated, specialisation pattern in medical chemistry and pharmacy and pharmacology. The results for the chemical industry confirm this. EU chemical firms do not use US-generated research to the same extent as pharmaceutical firms. Their home country knowledge base is relatively more specialised, in a persistent and integrated manner, in those fields which are particularly relevant to the innovative efforts of the chemical industry. Thus they rely heavily on the public research of their own country or other European countries.

Particular attention should be devoted to specialisation by type of research in EU countries. It was noted above that they are positively specialised in either medical chemistry or pharmacology at the level of applied and developmental research. However, those two fields do not show up as areas of positive specialisation in basic research (Table 33.5). Also, when we regressed the symmetric RSI in 1996 on the 1989 value for basic research alone, only the US, Japan, and the UK have β coefficients with a significance level higher than 5 per cent respectively, 1.04 (1 per cent), 0.95 (2 per cent), and 0.44 (4 per cent). The US, with both $\beta > 1$ and $\beta/R > 1$, increased specialisation in sectors where it was already specialised, and became less specialised where initially specialisation was low. Japan, with both $\beta \sim 1$ and $\beta/R \sim 1$, showed a high stability in its specialisation patterns. In particular, the US deepened its specialisation in fields related to the pharmaceutical industry: medical chemistry (C6) and pharmacology (C10). The four largest European countries saw an increase in the dispersion of their basic research specialisation. EU countries, especially Germany and France, show a tendency to remain more focused on traditional chemistry fields. No statistically significant results were obtained in the cases of applied research and applied technology and engineering. Therefore these data do not allow us to talk about a 'European paradox', according to which EU firms would not be capable of exploiting an efficient basic research system because of lack of 'development' capabilities. Our interpretative framework and data seem to point to the fact of these types of capabilities existing. What is missing is the basic research bit, with the result that EU pharmaceutical firms have to source research results from the US. The pattern of sourcing is consistently different when chemical firms are considered, as their home country knowledge bases seem more capable of providing basic research capabilities.

Despite the limitations of the data and the simplicity of this analysis, the location of different countries along the grid defined by the measures of persistence and integration both matches with a few things we know about the institutional structure of each country and also raises some interesting questions. For instance, the results concerning the 15 countries of the EU as a whole are hardly surprising. An EU–wide system of innovation is still in the process of formation. National industry and science and technology (S&T) policies still heavily influence country level specialisation patterns, preventing them from converging toward a homogeneous whole.

4. CONCLUSIONS

This chapter aimed at proposing data sources and methods of analysis to simultaneously look at the related processes of specialisation and integration.

Many have stressed that new knowledge is generated via a process of progressive specialisation, with new fields of knowledge developing out of pre-existing fields, and quite often complementing them, rather than replacing them within the portfolio of competence held by firms and countries. Hence, due attention is devoted to the analysis of patterns of specialisation in scientific and technological fields (through patent and bibliometric analysis) and the stability and persistence of such patterns over time. This paper has built upon a wealth of research in the strategy and innovation management literature to argue that the counterpart of specialisation is integration. The key competitive skill of firms (and countries) in modern economies lie in their ability to integrate into a coherent whole skills, competence and physical components that are developed within dense networks of heterogeneous organisations. Such integrative skills have received plenty of attention in the management literature, but very little effort has been devoted to the development of indicators based upon publicly available data that can, in principle, be used to go beyond the specificities of single firms or sectors. This chapter has first discussed the concept of 'coherence', as developed in Nesta (2001), and has then introduced 'depth' as a possible indicator of these integrative capabilities.

Much more needs to be done. First, a number of limitations of the data used in this paper must be borne in mind. For example, it would be of paramount importance to extend this analysis to citations to patent literature. In addition, the citations we used are those reported in the patent itself. However, only a subset of these citations was selected by the inventor. Others were added by the examiner, who might also have eliminated some of those chosen by the inventor. Are there systematic differences between the citation patterns of inventors, and those patterns resulting from the analysis of the overall list of citations found in the patent? Moreover, longer time series are needed to strengthen the results. Long citation lags may exist between publication dates and the time when those publications become recognised and used (and thus cited). Such citation lags may vary across disciplines or types of research. These issues need to be explored by developing wider data bases. Longer time series would also allow sensitivity analysis over the length of the sub-periods analysed in this paper. Moreover, this chapter has focused entirely upon the pharmaceutical industry. Cross-industry comparisons are needed in order to assess whether the proposed indicators capture some key characteristics of firms' technological specialisation strategies, or some inherent properties of their knowledge bases. Another issue to be considered is that in this chapter we have consolidated corporate groups in 1997. That is to say, we have made no distinction between capabilities acquired through internal development, and capabilities obtained through acquisitions. Given the frequency of M&As in the pharmaceutical industry, this is a key limitation in interpreting the indicator of depth as a measure of groups' integrative capabilities.

In addition, this type of analysis needs to be furthered in several directions, one of which is particularly needed. The indicators developed here (and particularly depth) need to be validated against indicators of innovative and economic performance. Can this indicator help to distinguish between innovative and non-innovative firms, and countries? Qualitative studies hint at a positive answer, but more research is needed to validate the specific indicator of integrating capabilities proposed in this chapter. This is a vital question to answer, if we want to maintain and strengthen the connection between 'indicators work' and theoretical and empirical developments in fields, such as strategy and innovation management, that look closely at the unfolding of processes of learning and innovation in 'real life' settings, and thus provide invaluable insights to researchers interested in developing empirically robust and theory informed indicators of science and technology.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge financial support from the Commission of the European Communities — EUROSTAT, IST project 'NewKInd', IST 1999-20728. Stefano Brusoni also acknowledges financial support from the ESRC (Economics and Social Sciences Research Council) through their Post Doctoral Fellowship scheme. The editors of this volume, one anonymous referee and Lionel Nesta provided very useful comments. As usual, all remaining mistakes are the authors' only responsibility.

REFERENCES

Pharmaceuticals Groups. What do patent citations to non-patent literature reveal? Forthcoming in *Economics of Innovation and New Technology*.

Brusoni, S., Geuna, A. (2003). An international comparison of sectoral knowledge bases: persistence and integration in the pharmaceutical industry. *Research Policy*, 32, 1897–1912.

Cantwell, J.A. (1989). *Technological innovation and multinational corporations*. Oxford: Basil Blackwell Ltd.

Conant, J.B., Nash, L.K. (1964). *Harvard case histories in experimental science*, vol. 1. Cambridge, MA: Harvard University Press.

Dalum, B., Laursen, K., Villumsen, G. (1998). Structural change in OECD export specialisation patterns: De-specialisation and "Stickiness". *International Review of Applied Economics*, 12, 447–67.

Fai, F. (2003). *Corporate technological competence and the evolution of technological diversification*. Cheltenham: Edward Elgar.

Godin, B. (1994). *The relationship between science and technology*, Unpublished DPhil Thesis, SPRU, University of Sussex, Brighton.

Granstrand, O., Patel, P., Pavitt, K. (1997). Multi-technology corporations: why they have "distributed" rather than "distinctive core" competence, *California Management Review*, 39, 8–25.

Henderson, R.M. (1994). The evolution of integrative capability: innovation in cardiovascular drug discovery. *Industrial and Corporate Change*, 3, 607–630.

Henderson, R.M., Cockburn, I. (1996). Scale, scope and spillovers: the determinants of research productivity in drug discovery. *Rand Journal of Economics*, 27, 32–59.

Iansiti, M., Clark, K. (1993). Integration and dynamic capability: evidence from product development in automobiles and mainframe computers. *Industrial and Corporate Change*, 3, 557–605.

Lawrence, P.R., Lorsch, J.W. (1967). *Organization and environment: managing differentiation and integration*. Homewood, IL: Irwin, 1967.

Loasby B. J. (1999). *Knowledge, institutions and evolution in economics*. London: Routledge.

March, J., Simon, H. (1958). *Organizations*. New York: Wiley.

Meliciani, V. (2001). *Technology, trade and growth in OECD countries. Does specialisation matter?* London: Routledge.

Nelson R.R., Winter, S. (1982). *An evolutionary theory of economic change*. Cambridge: The Belknap Press, Cambrdige.

Nesta, L. (2001). *The coherence of knowledge bases and technical change. Evidence from biotechnology firms between 1981 and 1997*, unpublished PhD thesis, University Pierre Mendes-France, Grenoble (in French).

Nesta L., Dibiagio, L. (2004). Knowledge organisation and firms' specialisation in the biotech industry. Forthcoming on *Industry and Innovation*.

Nesta L., Saviotti, P. (2003). *Intangible assets and market value: Evidence from biotechnology firms*. Paper presented at EMAEE 2003, University of Augsburg (D), April 10–12.

Pavitt, K. (1989). *International patterns of technological accumulation*. In N. Hood and J.E. Vahlne (Eds.), Strategies in Global Competition. London: Croom Helm.

Pavitt, K. (1998). Technologies, products and organization in the innovating firm: what Adam Smith tells us that Schumpeter doesn't. *Industrial and Corporate Change*, 7, 433–452.

Perrow, C. (1967). A framework for the comparative analysis of organizations. *American Sociological Review*, 32, 194–208.

Pisano, G.P. (1997). *The development factory: unlocking the potential of process innovation*. Boston MA: Harvard Business School Press.

Powell W., Owen-Smith J., Pammolli F., Riccaboni, M. (2002). A comparison of U.S. and European University-Industry relations in the Life Sciences. *Management Science*, 48, 24–43.

De Solla Price, D.J. (1963). *Little Science, Big Science*. New York: Columbia University Press.

Prencipe, A. (1997). Technological competencies and products evolutionary dynamics: A case study from the aero engine industry. *Research Policy*, 25, 1261–1276.

Prencipe, A. (2000). Breadth and depth of technological capabilities in CoPS: The case of the aircraft engine control system. *Research Policy*, 29, 895–911.
Soete, L.L.G. (1987). The impact of technological innovation on international trade patterns: The evidence reconsidered. *Research Policy*, 16, 101–130.
Rumelt, R.P. (1974). *Strategy, Structure, and Economic Performance*. Boston: Harvard University Press.
Teece D.J., Rumelt, R.P., Winter, S., Dosi, G. (1994). Understanding corporate coherence: theory and evidence. *Journal of Economic Behavior and Organization*, 23, 1–30.
Wang, Q., von Tunzelmann, G.N. (2000). Complexity and the functions of the firm: breadth and depth. *Research Policy*, 29, 805–818.

APPENDIX

SCI codes in fields related to chemistry and life sciences

Code	*Discipline*	*Code*	*Discipline*
AD	agriculture, dairy and animal sci	JY	food science and technology
AH	agriculture, multidisciplinary	KI	gastroenterology and hepatology
AM	agronomy	KM	genetics and heredity
AQ	allergy	MA	hematology
AY	anatomy and morphology	MU	horticulture
AZ	andrology	NE	public, environmental and occupational health
BA	anesthetology	NI	immunology
CN	behavioral sciences	NN	infectious disease
CO	biochemical research methods	OP	medicine,legal
CQ	biochemistry and molecular biology	PW	medical laboratory technology
CU	Biology	QA	medicine,research & experimental
CX	biology, miscellaneous	QB	medicine, miscellaneous
DA	Biophysics	QU	microbiology
DB	biotech and applied microbiology	RO	multidisciplinary sciences
DE	plant sciences	RQ	mycology
DM	Oncology	RT	clinical neurology
DQ	cardiac and cardiovascular systems	RU	neurosciences
DR	cell biology	SA	nutrition and dietetics
DS	critical care medicine	SD	obstetrics and gynaecology
DW	chemistry, applied	TI	parasitology
DX	chemistry medicinal	TM	pathology
DY	chemistry, multidiscinary	TQ	pediatrics
EA	chemistry, analytical	TU	pharmacology and pharmacy
EC	chemistry, inorganic and nuclear	UM	physiology
EE	chemistry, organic	UY	polymer science

Code	*Discipline*	*Code*	*Discipline*
EI	chemistry, physical	WE	respiratory system
FF	emergency medicine	WH	rheumatology
FQ	cellular biology and histology	YA	surgery
FY	dentistry, oral surgery and medicine	YO	toxicology
GA	dermatology and venereal disease	YP	transplantation
HY	developmental biology	YU	tropical medicine
IA	endocrinology and metabolism	ZA	urology and nephrology
II	engineering, chemical	ZC	veterinary sciences
IQ	engineering, electric and electronic	ZD	peripheral vascular disease
IY	entomology	ZE	virology

Chapter 34

SCIENCE AND TECHNOLOGY SYSTEMS IN LESS DEVELOPED COUNTRIES

Identifying a Threshold Level and Focusing in the Cases of India and Brazil

Eduardo da Motta e Albuquerque

Universidade Federal de Minas Gerais, CEDEPLAR, Belo Horizonte, Brazil
E-mail: albuquer@cedeplar.ufmg.br

Abstract: Patent and scientific publication data offer useful information for the analysis of key features of national systems of innovation of less developed countries. However, the use of these statistics should be subjected to careful evaluation. This chapter uses data for 120 countries (patents, scientific papers, and GDP), investigating different patterns of interactions between science and technology according to different levels of development. Later this chapter focuses on the data for India and Brazil, investigating interactions between science and technology in three dimensions: inter-sectoral, inter-regional, and inter-temporal.

1. INTRODUCTION

As discussed in previous chapters of this Handbook, developed countries have systematic activities in science and technology and systematic and well organised data on these activities. This is not the case of a broad set of countries identified as 'less developed countries' (LDCs henceforth).

In the case of LDCs the difficulties to find reliable data are not trivial. These difficulties have pushed scholars as Coe, Helpman and Hoffmaister (1995) in their study on "North–South" technological diffusion to use "trade openness" as a proxy for access to international R&D: they have argued that there are no reliable data on R&D for LDCs. Furthermore, broad samples as those provided by the Penn World Table do not use indicators of science and technology (Barro et al., 1995).

H.F. Moed et al. (eds.), Handbook of Quantitative Science and Technology Research, 759-778.

This chapter suggests ways of dealing with S&T data for LDCs. At least two indicators are under-utilised: statistics from the United States Patent and Trademark Office (USPTO henceforth) patents and from the *Institute for Scientific Information* (ISI, henceforth) scientific papers. Both statistics are available on the Internet (www.uspto.gov and www.isiknowledge.com) and researchers may prepare databases suitable for multifarious research purposes. Naturally, these indicators have pros and cons. The pros and cons of patent and scientific paper statistics have been broadly discussed (Griliches, 1990; Patel and Pavitt, 1995; Grupp, 1998). Notwithstanding, for less developed countries more problems must be considered: different levels of development, and different levels of maturity of national systems of innovation (NSIs henceforth) are reflected in the ability of these statistics to capture all innovations relevant for a given country in a (relatively) more backward level of development. The specificities of these statistics for LDCs are discussed in the next section.

Why have these data been under-utilised? Probably because there is a 'mantra': statistics of patents and scientific papers are neither relevant nor useful for investigations concerning LDCs. This chapter shows that, on the contrary, patents and scientific papers statistics are relevant for LDCs. Moreover, these statistics have been powerful indicators of successful catching up processes.

Another problem concerning studies of LDCs is the procedure to group all LDCs as if they were all the same. Although ordinary rankings prepared by the World Bank or by the United Nations Development Programme disaggregate LDCs in different categories (low and middle income countries, and lower middle and upper middle income countries), it is not unusual for researchers to discard these differences and to deal with LDCs as if they were the same thing. S&T data may contribute to avoiding this inaccuracy. Science and technology data, for instance, help to differentiate India and China from the other 'low income countries' in the World Bank rankings (World Bank, 1997).

This chapter suggests a preliminary differentiation of LDCs using USPTO patents and ISI indexed papers statistics. This suggestion may be seen as a dialogue with Amsden's (2001) elaboration. Amsden divides the LDCs into two broad sets: the 'rest' and the 'remainder'. It is an interesting division, but it could be improved. In the 'rest' Amsden ranks South Korea, Taiwan, India, Mexico and Brazil, amongst others. However, as catching up countries, South Korea and Taiwan probably have left the 'rest' during the 1980s and the 1990s. Indeed, the data on patents and scientific papers may help to differentiate the trajectories of catching up countries and the 'rest'.

This chapter later focuses on two countries (India and Brazil) and investigates them using USPTO patents and ISI indexed papers.

The main objective of this chapter is a double invitation: first, to a broader use of these under-utilised (and rich) sources of statistical data; and second, to further and co-operative research to improve the data and the knowledge of the scientific and technological situation of LDCs.

2. METHODOLOGICAL REMARKS: ADVANTAGES AND DISADVANTAGES OF SCIENCE AND TECHNOLOGY STATISTICS FOR LDCs

Taking as the starting point the two sets of statistics discussed in the initial sections of this Handbook (patents granted by the USPTO and papers indexed by the ISI) it is necessary to stress the specific advantages and disadvantages of these data for LDCs.

The advantages are clear: 1) data easily collected, as they are available at the Internet; 2) it is possible to build long time series; 3) they are internationally comparable data, because firms, individuals, and institutions of different countries must follow the same rules to apply for a patent at the USPTO or to publish a paper indexed by the ISI; 4) they are reliable data.

Statistics of USPTO patents and ISI indexed papers have advantages in regard to others indicators of science and technology. R&D indicators are not systematically collected in all LDCs[1]. Scientists and engineers per inhabitants are not easily available[2]. Innovation surveys are limited to a few countries. Therefore, although imperfect, data on patents and papers have advantages vis-à-vis other indicators.

Papers are not a perfect measure of scientific production, and patents are not a perfect measure of technological innovation. The literature has both used these data and warned about their problems, limitations and shortcomings (see previous chapters in this Handbook).

Scientific papers, the data collected by the ISI, have various shortcomings, from language bias to the quality of research performed: there could be important research for local needs that does not translate into international papers, but only in national publications not captured by the ISI database. There is a huge literature on the problems of this indicator (Patel

[1] In the Brazilian case only in 2002 was reliable data available for national R&D expenditures for the manufacturing sector. However, data for the service sector is not available (Viotti & Macedo, 2003).

[2] A look at the World Development Indicators (World Bank, 2003) shows that these data are not available to all countries in all years.

and Pavitt, 1995; Velho, 1987). Paper citations improve the quality of this indicator, but it would not be so useful for this paper, further biasing the data against papers produced in countries with little developed scientific institutions[3].

Patents, the USPTO data, also have important shortcomings, from commercial linkages with the US to the quality of the patent: again, local innovation necessarily is limited to imitation in the initial phases of development, and imitation or minor adaptations do not qualify for a patent in the USPTO.

In general the main shortcoming of USPTO patents and ISI indexed papers, as statistics of S&T for LDCs, is their feature of 'tips of the iceberg'. They do not represent the whole scientific and technological production of these countries.

For patents, as discussed in a previous paper (Albuquerque, 2000) on Brazil, there are important differences between patenting at the national office and at the USPTO. USPTO patents may not capture all innovations relevant for LDC. For instance, LDC may generate incremental innovations relevant for their absorptive capability (and for their process of development) without the originality necessary for a patent application. Therefore important data are not captured by these statistics. For instance, in the Brazilian case the steel industry is amongst the leading sectors at the national office but it disappears in the USPTO statistics. Another important difference is the position of research institutions: for the 1990s there are five of them amongst the top 20 at the national office (three universities, a health research institute, and an agricultural research institute) and none at the USPTO. This problem has also been identified for the Mexican case: the leading patent institution at the national patent office (between 1980 and 2002) is the *Instituto Mexicano Del Petrôleo*, which ranks only in the 25th position at the USPTO in the same period.

For scientific papers the well known language bias impacts strongly the ISI statistics. Sandelin and Sarafoglou (2003) show that there is a consistent bias in favour of English speaking scholars (who would write 'almost exclusively in English') and that scholars from non-English speaking countries tend to publish 'less in English and more in their domestic

[3] It is justifiable to study less developed countries with data from scientific papers because the existence of a scientific infrastructure hints at: 1) the level of development of the educational resources of the country; 2) the quality of their universities; 3) their connections with the international flows of scientific knowledge; and 4) the commitment of these universities with research activities. This assumption implies that the number of published papers may be taken as an indicator of the general situation of the educational conditions of the country and of their usefulness to the economic development.

language the larger is their domestic language'. Sandelin and Sarafoglou (2003) call for caution in using ISI databases for international comparisons. Therefore behind the papers indexed by the ISI there might exist a broader domestic production of scientific papers important for local activity.

Assuming that USPTO patents and ISI scientific papers statistics are 'tips of the iceberg', it is important to evaluate the contributions of domestic statistics.

First, data from National Patent Offices. Domestic patents of developing countries provide a better 'picture' of technological activities than USPTO patents. For instance, for the period between 1980 and 1995 there were 8,309 domestic resident patents and 475 USPTO patents granted to Brazilian residents, almost a 20:1 ratio (for the Indian case, see Rajeswari, 1996; for the Mexican case, see Aboites, 1996). Domestic patents will be selected and the best might have an application submitted to the USPTO. This selection mechanism may provide useful information.

But even the data of National Patent Offices may not capture all innovations. Innovative modifications may be made to foreign technologies, which may be copied or adapted to suit local patterns. These minor improvements, although locally relevant, are not straightforwardly translated into patents. Local learning may exist without local patenting (probably data on trade marks could be useful for informing about activities that mean some kind of innovation, relevant to the local economy but not patentable, trade marks may inform about activities related to product differentiation, which in its turn could signal the beginning of a local technological ladder towards more innovative activities).

When dealing with domestic patents, there are important statistical implications of different patent laws. Until recently, less developed countries (Brazil, India, and Argentina are good examples) forbade patents in sectors such as pharmaceuticals and food. Differences in bureaucratic procedures may lead to differences in patenting activities.

And for developing countries important technological improvements lie in transfer mechanisms (capital goods imports, technology licensing, etc) which, again, are not captured by patent statistics.

One important remark is about a limitation of patent statistics in relation to software technology: software has been a relevant product of India (D'Costa, 2002) and Brazil (MIT/SOFTEX, 2002) but its performance is not captured by these statistics.

For scientific papers data for domestic production is important. In the Brazilian case there are indications that the publications not indexed by the ISI are at least two times greater than the ISI papers of Brazilian authors. The scientific disciplines can be organised according to their different levels of internationalisation. But the main problem is the limited scope of these databases (see, for instance, www.scielo.org.br). In the Brazilian case, for instance, the database for national publications is new and has fewer papers than the Brazilian papers indexed by the ISI.

Thus there are problems with the USPTO and ISI data, but there are also problems with the domestic patents and papers databases. The most appropriate way to deal with these statistics would be a combined evaluation of data from the USPTO and from different National Patent Offices, but the enterprise requests international collaboration (and probably is the most important point in the agenda for further research of this chapter).

Therefore this paper acknowledges these important limitations, and this literature must be kept in mind to qualify the results discussed in the next sub-sections. Specially the 'tip of the iceberg' nature of the USPTO and of the ISI databases. Furthermore, despite these problems these two databases do provide useful and under-utilised information for research.

3. USING PATENT AND SCIENTIFIC PUBLICATION DATA TO LOCATE THE INTERNATIONAL POSITION OF SELECTED LDCS

To locate the international position of India and Brazil this section uses data for 120 countries (patents, scientific papers, and GDP) and is based on previous work (Bernardes and Albuquerque, 2003).

This section shows a threshold between the mature NSIs and the rest in regard to scientific and technological production. Looking through time these data show the movement of catching up countries away from relatively backward positions beyond the threshold level. The remarkable movement of catching up countries can be compared to a relatively stagnant position of South Africa, India, Mexico, and Brazil (section IV).

For this purpose data about GNP per capita (US dollars, PPP, according to the World Bank, for 1998), patents (for 1998, 1990, 1982 and 1974, according to the USPTO, 2001), and scientific papers (for 1998, 1990, 1982 and 1974, according to the *Institute for Scientific Information*, 2001) were collected for 120 countries[4].

[4] The countries are: Albania; Algeria; Argentina; Armenia; Australia; Austria; Azerbaijan; Belarus; Belgium; Bolivia; Bosnia and Herzegovina; Brazil; Bulgaria; Cameroon; Canada; Chile; China; Colombia; Congo (Dem. Rep.); Congo (Rep.); Croatia; Cuba; Czech Republic; Denmark; Dominican Republic; Ecuador; Egypt; El Salvador; Estonia; Ethiopia; Finland; France; Germany; Ghana; Greece; Guinea; Haiti; Honduras; Hong Kong

The range and usefulness of these indicators should be highlighted: there are 115 countries out of 120 which have published at least one scientific paper in 1998; and 89 countries out of 120 applied at least one patent at the USPTO in 1998. Only one country (Trinidad Tobago) out of 120 has zero patents and zero papers in 1998.

3.1 A Simple Model of Stages of Development, Science Production, and Interactions between Science and Technology

To perform the statistical analysis (in the next sub-sections) this sub-section puts forward a very simple model. This model describes the relationship and the interactions between science, technology and economic growth. It simplifies the complex and multifarious connections, interactions, and causal chains that constitute the province of economic growth. However, this model contributes to organising the data in a very simple way, differentiating countries between those which already produce science and technology, according to the proxies, and those which do not have both productions.

The theoretical background and the intuitions of this very simple model are found in the literature on the interactions between science and technology (for a review of this literature see Bernardes and Albuquerque, 2003, sections I, II, and III).

Three stylised facts can be drawn: 1) developed countries have strong scientific and technological capabilities, and there are interactions and mutual feedbacks between the two dimensions; 2) the rôle of science during the catching up process is crucial and it is two-fold: source of absorptive capability, and provider of public knowledge for the productive sector; 3) less developed countries are caught in a 'low-growth trap' given, *inter alia*, the low levels of scientific production.

(China); Hungary; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Jamaica; Japan; Jordan; Kazakhstan; Kenya; Korea (Rep.); Korea (Dem. Rep.); Kuwait; Kyrgyzstan; Latvia; Lebanon; Lesotho; Libya; Lithuania; Macedonia; Madagascar; Malaysia; Malawi; Mali; Mauritania; Mauritius; Mexico; Mongolia; Morocco; Myanmar; Namibia; Nepal; Netherlands; New Zealand; Niger; Nigeria; Norway; Oman; Pakistan; Panama; Paraguay; Peru; Philippines; Poland; Portugal; Romania; Russia; Saudi Arabia; Senegal; Sierra Leone; Singapore; Slovakia; Slovenia; South Africa; Spain; Sri Lanka; Sudan; Sweden; Switzerland; Taiwan; Tanzania; Thailand; Trinidad and Tobago; Tunisia; Turkey; UK; USA; Uganda; Ukraine; United Arabic Emirates; Uruguay; Uzbekistan; Venezuela; Vietnam; Yemen; Yugoslavia; Zambia; and Zimbabwe.

To suggest this very simple model six steps are necessary: 1) the first step is the recognition of two different dimensions of innovation–related activities (the scientific infrastructure and the technological production); 2) the second step is the identification of the division of labour between them; 3) the third step is the identification of interactions between the scientific and technological dimensions, as well as the dynamics of these interactions; 4) the fourth step is the suggestion that these interactions change during the development process, reaching at last a level of strong and mutual reinforcing relationships found in developed economies; 5) the fifth step is the conjecture that this evolutionary path depends on the scientific infrastructure (at least, the improvement and the growth of the scientific infrastructure is a *necessary* but not sufficient condition for initiating technological development), and that there are thresholds of scientific production which must be overcome to reach new stages (and new levels of interaction between science and technology); 6) finally, these interactions in the science and technology field might be integrated in the causal chains of economic growth.

The data gathered for this paper provide one feature for this simple model: amongst the 115 countries which produced at least one paper in 1998 85 countries were granted at least one patent. The 30 countries with scientific production but no patents are the countries that compose a special class: countries in a 'low growth trap', where their scientific production is so low that it does not yet feed technological production (these countries are included in 'Régime I', below).

These steps and comments lead to the very simple model: there are three different 'régimes', ranging from the least developed countries ('régime' I) to the developed countries ('régime' III).[5]

The very simple model suggests that as the 'régimes' change, the number and the channels of interactions between scientific infrastructure, technological production and economic growth concomitantly also change. As the country evolves, more connections are 'turned on' and more interactions operate. The 'régime' III is the case in which all connections and interactions are working (they have been 'turned on' during previous phases).

As long as the development takes place, the rôle of 'others' in the causation of economic growth decreases. In other words, as a country

[5] The term 'régime' is not a good one, but it is useful to delimit the different forms of operation of the relationship and interactions amongst the four variables used in the model in its present (and very initial) level of elaboration. The three régimes are used as synonyms for 'mature' NSIs (Régime III), 'immature' NSIs (Régime II) and countries without or with very weak S&T institutions (Régime I).

upgrades its economic position, its economic growth is more and more 'caused' by its scientific and technological resources. The mutual feedbacks between them contribute to explaining why the modern economic growth is fuelled by strong scientific and technological capabilities (Fagerberg, 1994).

This very simple model is suggested in order to enable the data analysis of next sections, focusing the interactions between science and technology.

3.2 Correlation between Scientific and Technological Production and GNP per Capita

There is a correlation between GNP per capita, number of articles per million of inhabitants (A* henceforth) and the number of patents per million of inhabitants (P* henceforth). The data are for the year 1998. Only countries with data available and scores different from zero are represented.

Table 34.1. Averages and standard deviation of articles per million inhabitants (A*); patents per million inhabitants (P*); and the ratio between articles per million inhabitants and patents per million inhabitants (A*/P*), according to the their income level (GNP per capita) in 1998

Group of Countries (GNP per capita)	*A**		*P**		*A* / P**	*Number of countries*
	Average	Standard deviation	Average	Standard deviation	Average	
> US$ 19,000	937.99	377.69	154.42	121.54	6.07	19
US$ 10,000–US$ 19,000	476.59	432.32	64.68	107.37	7.41	13
US$ 5,000–US$ 10,000	115.68	133.58	1.45[a]	1.76	79.78	25
US$ 3,000–US$ 5,000	40.87	50.10	0.43[b]	0.58	95.04	17
< US$ 3,000	14.79	25.06	0.10[c]	0.18	147.90	40
GNP not available	14.81	28.89	0.04[d]	0.10	370.25	6

Source: World Bank, 2000; USPTO, 2001; ISI, 2001 (authors' elaboration).
[a] 3 countries (with P* = 0). [b] 2 countries (with P* = 0). [c] 21 countries (with P* = 0).
[d] 5 countries (with P* = 0).

Table 34.1 organises the data (patents per million inhabitants, scientific papers per million inhabitants and a ratio between these two data) according to countries income levels.

Table 34.1 shows the correlation between science, technology and income, as the scientific and technological production are directly related to the income level. The scientific and technological productions are higher for the richer countries (for GNP per capita greater than US$ 19,000, A* = 937.99; P* = 154.42) than for poorer countries (for GNP per capita less than US$ 3,000, A* = 14.79, P* = 0.10).

Table 34.1 presents an initial hint about the existence of thresholds of scientific production. The third column presents the ration between A* and P* (the ratio A*/P* is calculated dividing the average A* and average P* for each group of countries). This ratio may be understood as an indicator of efficiency in the transformation of scientific production into technological outputs. The more efficient a group of countries is, the smaller is the ratio (the countries in that group, in average, produce more patents for a given stock of scientific papers).

In addition, one remark is necessary. Countries with zero patents or zero scientific papers have been excluded from Table 34.1 (115 countries out of 120 have published at least one scientific paper in 1998; and 89 countries out of 120 applied at least one patent at the USPTO in 1998). There are 30 countries with scientific publications but without USPTO patent, which constitute the ‘régime’ I. These 30 countries have not reached even the first threshold, the threshold necessary to trigger the beginnings of a technological production (as captured by the proxy of USPTO patents).

The next step in this analysis is to divide the sample countries between the three régimes suggested in the previous sub-section. So far it is only possible to indicate the general relationship between income, science and technology, and to identify the countries included in the “régime I” (the 30 countries with papers but no patents). To divide between ‘régimes’ II and III it is necessary to investigate thresholds of scientific production.

3.3 Preliminary Evidences about Thresholds of Scientific Production

Table 34.1 suggests the existence of two behaviours in the relation between A* and P*. The remainder of this section discusses and presents preliminary statistical evidences about the existence of thresholds between different stages of development.

3.3.1 The threshold in 1998 data

The crossover and the threshold level can be better observed in Figure 34.1. Figure 34.1 displays the data for the year 1998 in a two-dimensional plot in log-log scale. In this plot it is possible to define two regions. Roughly speaking, they are separated by the point (A* ≈ 100 and P* ≈ 1). The technologically immature countries are at the left or lower than this point and the mature countries at right/upper.

In this Figure countries representative of different stages of developed are pinpointed: 1) Régime III countries as USA, Japan and Switzerland; 2) catching up countries, now part of Régime III, such as South Korea and Taiwan; 3) Régime II countries, as South Africa, India, Mexico and Brazil (countries that will be focused in section III, below).

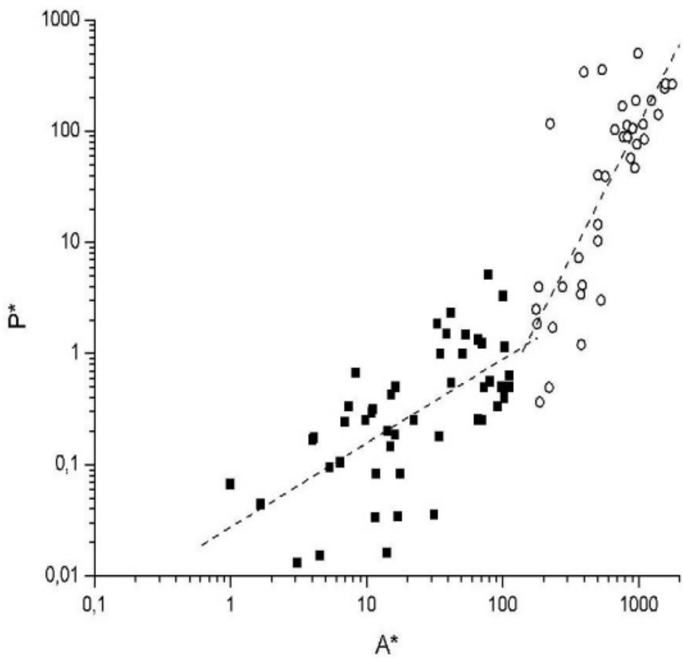


Figure 34.1. Log-log plot of articles per million of inhabitants versus patents per million of inhabitants for the year 1998. Here the two subsets are identified by different symbols. Two power functions have been used to fit the two subsets.
Source: Bernardes and Albuquerque (2003).

Those points can be fitted by two power functions $P^* \propto (A^*)^{\beta}$, which has been done by dividing the set of points in two subsets, which are shown by different symbols (filled squares and open circles) in Figure 34.1.

The crossover between the two lines occurs at A* ≈ 150. This is the threshold that identifies the transition from 'régime' II to 'régime' III.

The data plotted in Figure 34.1 give important clues for the behaviour of the interactions within the process of development: these data suggest a non-linear dynamics. In contrast to different approaches that assume linear relations between science (or scientific production), technology (or technological production) and economic growth, with one variable determining the other in a unidirectional chain of causation.

The threshold is not observed only in the year 1998. The same behaviour can be observed at different times. One interesting aspect is that the value of this threshold seems to double from one period to another: in 1974 the threshold was 7 A*, in 1982 28 A*, in 1990 60 A* and in 1998 150 A*. This moving threshold could be interpreted as a signal of the increasing role of science in newer technological paradigms, supporting empirically Dosi's suggestion (1988, p. 1136) and OECD's report (2002, Chapter 1). Additionally, as a corollary, this indicates the inter-temporal increase in the weight of the scientific infrastructure as a precondition for the beginning of a catching up process.

3.3.2 The three régimes: general overview

With the thresholds identified, it is now possible to resume the analysis from sub-section 3.2, focusing on how the performance of the sample varies according to the three régimes.

Table 34.2 reorganises the data, distributing the 115 countries according to their 'régimes' in 1998.

Table 34.2. Averages and standard deviation of articles per million inhabitants (A*); patents per million inhabitants (P*); and the ratio between articles per million inhabitants and patents per million inhabitants (A*/P*), according to the their 'régime' (Figure 34.1) in 1998

	*A**		*P**		*A*/P**		*PPP GNP per capita*	
'Régime'	Ave-rage	St-dev	Ave-rage	St-dev	Ave-rage	St-dev	Ave-rage	St-dev
III (n = 38)	666	419	94.91	118.95	71	122	16,698	7,008
II (n = 47)	38	35	0.65	0.94	144	191	4,431	2,626
I (n = 30)	12	17	0	0	-	-	1,635	1,443

Source: World Bank, 2000; USPTO, 2001; ISI, 2001 (authors' elaboration)

Table 34.2 highlights features of these different 'régimes', but also presents elements that call for caution in the analysis.

According to Table 34.2, as countries evolve from régime I to régime III the averages of all indicators increase (scientific production, technological production and income). It is interesting to note that the ratio A*/P* decreases as the scientific production increases: as the scientific production grows, the capacity of the technological sector to use this knowledge increases, becoming more efficient in the transformation of scientific information into technological products.

Probably this means that at the ‘régime’ III, there are more connections ‘turned on’ and more interactions working. Probably mutual feedbacks and virtuous cycles are working. On the other hand, Tables 34.1 and 34.2 show that as the income level falls, the efficiency of the transformation of scientific production into technological output also falls (the ratio A*/P* increases according to Table 34.2’s fourth column). In other words, probably there are fewer connections, fewer and weaker interactions, unidirectional causal links, making room for low growth traps: the cases of ‘régimes’ I and II take place.

Table 34.2 also highlights some limitations of this analysis. Table 34.2 shows that for ‘régimes’ II and III the averages for P* and A*/P* are smaller than the standard deviation, showing a large variance within these two set of countries. Table 34.2 also reveals that for ‘régime’ I the average for A* is smaller than the standard deviation.

These data hint that the interactions between science and technology seem to be triggered after a certain threshold of scientific production has been attained. Or in a more cautious statement: the attainment of a threshold of scientific production seems to be a precondition for improved technological production.

4. FOCUSING SELECTED LDCs: THE CASES OF INDIA AND BRAZIL

Once the international position of India and Brazil has been shown (Figure 34.1, below the threshold level; Table 34.2, ‘régime’ II), this section focuses in their internal data, attempting to point differences among their science and technology systems. A closer look might be informative. The main objective of this section is the investigation of interactions between science and technology.

4.1 Inter-Sectoral Interactions between Science and Technology

Table 34.3 aggregates the USPTO patents granted to India and Brazil according the technological sub-domains of the classification suggested by the *Observatoire des Sciences et Téchniques* (OST henceforth, 2000, p. 409). Table 34.3 shows differences in the leading technological domains in India and Brazil.

Table 34.3. Leading technological domains of USPTO patents, according to the OST classification (1981–2001)

Country	*OST technological sub-domain*	*Patents*
India	Organic Chemicals	194
	Pharmaceutical-Cosmetics	146
	Basic Chemicals	70
	Macromolecular Chemistry	49
	Technical Procedures	46
	Biotechnology	44
	Informatics	42
	Materials-Metallurgy	37
	Total	883
Brazil	Household Consumption	85
	Medical Engineering	80
	Mechanic Components	77
	Construction	75
	Maintenance-Printing	73
	Motors-Pumps-Engines	71
	Technical Procedures	65
	Total	1,172

Source: OST, 2003; USPTO, 2001 (author's elaboration)

India has a broader presence of more R&D–intensive areas vis-à-vis the Brazilian case. These data are coherent with data collected (for 1997) by the OST (2000, p. 325), where India presents a specialisation in 'fine chemicals and pharmaceutics' and 'basic chemicals and metallurgy', while Latin America presents a specialisation in 'basic chemicals and metallurgy', 'equipments, mechanics, transport' and in 'household consumption'. The presence of biotechnology and informatics between the leading technological domains in the India case should be noted.

Table 34.4. Scientific Revealed Comparative Advantage (SRCA): immature NSIs for the year 2001

Country	*Discipline*	*SRCA*
INDIA	Agriculture/Agronomy	5.467
	Biotechnol & Appl Microbiol	3.390
	Veterinary Med/Animal Health	3.193
	Organic Chem/Polymer Sci	2.789
	Multidisciplinary	2.669
	Engineering Mgmt/General	2.498
	Metallurgy	2.481
	Chemistry	2.471
	Food Science/Nutrition	2.386
	Materials Sci and Engn	2.247
	Chemical Engineering	2.014
BRAZIL	Agriculture/Agronomy	3.976
	Dentistry/Oral Surgery & Med	3.234
	Biology	2.761
	Entomology/Pest Control	2.482
	Biotechnol & Appl Microbiol	2.196
	Medical Res, General Topics	2.167

Source: ISI, 2003 (author's elaboration)

Table 34.4 aggregates the data of scientific publications according to ISI sub-disciplines. Following a suggestion from Pavitt (1998), this section uses an indicator proposed by Lattimore and Revesz (1996): the Scientific Revealed Comparative Advantage (SRCA henceforth)[6]. The scientific specialisation of India and Brazil has in common Agriculture/Agronomy as leading disciplines. This leading position hints a focus of science on domestic needs[7].

India and Brazil probably fit in the 'mixed' case, according to the 'pattern of comparative advantage in publications' suggested by Lattimore and Revesz (1996, p. 14). Table 34.4 shows India with leading disciplines that are 'natural resources based', 'medical' and 'industry based'. Brazil has disciplines 'natural resources based' and 'medical'.

A comparison between Tables 34.3 and 34.4 may be carefully done, as this comparison might provide hints on the inter-sectoral interactions of science and technology.

[6] SRCA = (Pi,j/Pi,world)/(Pallfields,j/Pallfields,world) (Lattimore & Revesz, 1996, p. 15), where P = scientific papers; from the country i, and scientific field j.

[7] This point was highlighted by the editors' comments on a previous version of this chapter. The data on South Africa show the leading position of the discipline Geology/Petrol/MiningEngineering.

India displays a correlation between the leading technological domains (Table 34.3) and the leading scientific disciplines (Table 34.4). Chemistry and related disciplines are in leading positions and Organic Chemicals leads the patent data, followed by related sectors (Pharmaceuticals and Basic Chemicals. The rise of Biotechnology in the scientific dimension and leading positions of other health–related disciplines (Veterinary, SRCA = 3.193) should be noticed.

Brazil has Agriculture/Agronomy in the leading position, as in India. Data from the Brazilian Patent Office rank EMBRAPA (the leading institution in agricultural research) in the 6^{th} position for the period between 1990 and 2000. The concentration in 2001 in biology and health–related disciplines might be feeding interactions with the health sector, which has an expressive presence in the patent statistics. The position of the health–related disciplines might also be related to the formation of new biotech firms.

4.2 Other Dimensions of Interactions between Science and Technology

4.2.1 Inter-regional interactions

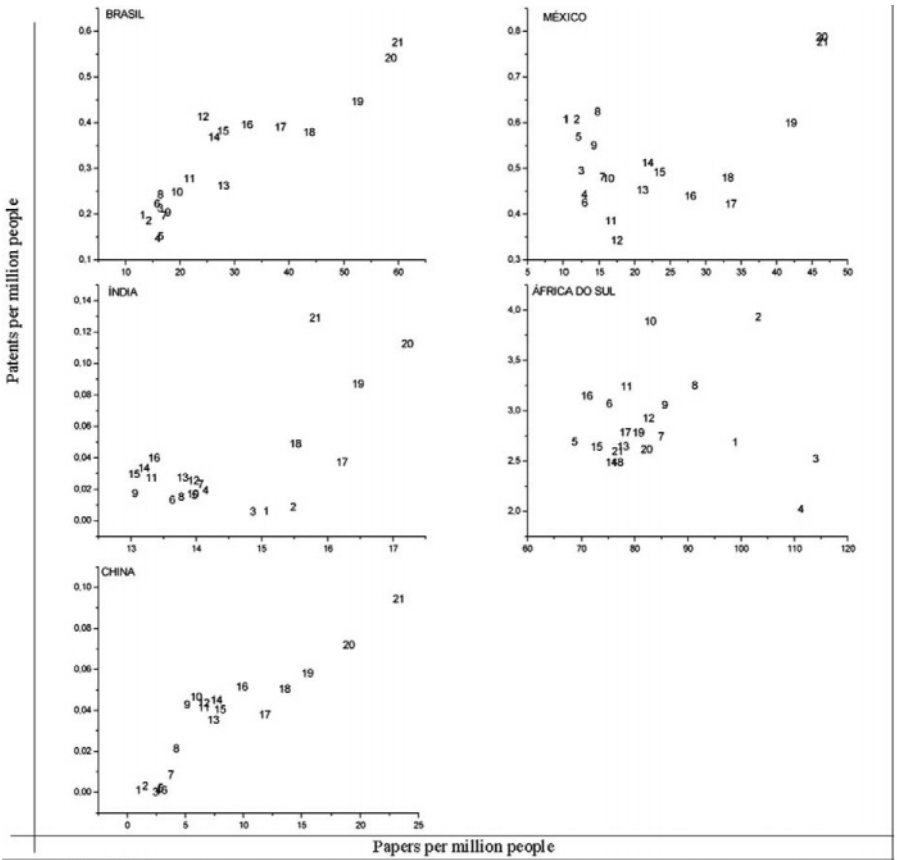
Inter-regionally the question is whether or not there is a correspondence of the leading region in technological production and the leading region in scientific production. To investigate this the data on patents and on papers are organised to capture the address of the patent-owner or the author of the paper.

The result shows that India and Brazil have the same state leading both the technological and the scientific production (Maharashtra and São Paulo, respectively).

4.2.2 Inter-temporal interactions

Inter-temporally the question is whether or not do the two dimensions co-evolve. Silva (2003) investigates this dimension, finding a sort of ‘polynomial relationship’ between the data for articles per million people and patents per million people for various developed countries and for catching up countries. Silva (2003) shows a non-linear relationship between improvements in the scientific dimension and in the technological dimension.

Silva (2003) organises data for ‘immature’ NSIs and the graphs shown in Figure 34.2 are drawn from his work. South Africa, Mexico, and China are included for comparative reasons. The observation of these inter-temporal trends may provide another important piece of information: an overall evaluation of the performance of these countries during two decades (which in the Latin American countries has been called as the 'lost decades'). Although hard economic times, in terms of the S&T dimension, the situation was not of pure decline. Figure 34.2 shows that for India, Mexico, and Brazil the last year of the time series (year 2000, dot 20, in the Graphs) is in a better position vis-à-vis the first year (year 1980, dot 1) of the time series (both in papers per million people and patents per million people). South Africa is the exception. Brazil seems to have resisted well, with a gradual rise in scientific and technological terms throughout all the period, although in relative terms the Brazilian share in the world technology is almost the same when 1980 is compared to 2000 (but this is a positive result).



Source: Silva (2003)

Figure 34.2. Patents per million people and papers per million people (selected "immature" NSIs)

According to Silva, among the 'immature' NSIs only Brazil displays the 'polynomial relationship' identified for developed and catching up countries. China also displays this pattern.

What do Brazil and China have in common, according to Figure 34.2? They show a constant increase in their scientific production. Presumably this is an important reason for a positive relationship between science and technology. In the Mexican case, from 1991 (dot 11) onwards the scientific production has resumed a consistent growth pattern and a 'polynomial pattern' can be seen.

With respect to the position of the scientific production in 1981 (see dot 1), from Figure 34.2 it can be seen that for South Africa and India this year's production is not the lower of the whole period. Thus for both South Africa and India at least a partial decline in scientific production took place, a general decline for the South African case, and partial decline with a further increase for the Indian case.

In the South African case the government reports a drop in R&D expenditures between 1990 (1.1% of the GDP) and 1994 (0.7% of the GDP) and the beginning of a structural rearrangement in the post-apartheid era (The Government of the Republic of South Africa, 2002, p. 15). This report mentions the "termination of key technology missions (such as military dominance in the subcontinent and energy self-sufficiency) by the previous government" (p. 15). Certainly there are huge costs in a transition to post-apartheid NSIs, with more people to serve and new needs to satisfy.

5. CONCLUDING REMARKS AND AN AGENDA FOR FURTHER RESEARCH

USPTO patents and ISI scientific papers statistics, although a 'tip of the iceberg' of domestic technological and scientific productions, are useful to differentiate and to cluster countries according to levels of development in a way that takes into account features of their scientific and technological positions. Using these data, the major points are:

1. Data from 120 countries have been used to disaggregate them in three different 'régimes' (these régimes may correspond to levels of formation of NSIs, that range from countries with 'mature' NSIs to countries without or with weak science and technology institutions), identifying a threshold level in terms of scientific production (150 articles per million people in 1998).
2. India and Brazil share an international position below the 'threshold level' of mutually reinforcing science and technology interactions, below

the 'critical mass' level for and adequate science and technology production.

3. India and Brazil have a 'mixed' pattern of comparative advantage in scientific publications, both presenting Agriculture/Agronomy in the leading position, although they have important differences in scientific specialisation of these countries.
4. India and Brazil have different technological specialisation; India is present in more R&D–intensive sectors than Brazil.
5. The investigation of interactions between science and technology in countries below the threshold level can be carried out by investigating three different levels of interaction: inter-sectoral, inter-regional, and inter-temporal.

The main suggestion of this chapter for further research is this three-dimensional investigation of interactions in LDCs (investigation that may be extended with selected case studies to inform more precisely the data analysis).

This investigation on India and Brazil shows that USPTO patents and ISI papers statistics are useful for LDCs. Naturally this is a very preliminary and tentative effort which needs improvement.

One key point in the agenda for further research is the use of National Patent Offices' patent data to complete and to compare with USPTO data. Another point is the enlargement of the data on scientific papers to include the domestic production of papers not indexed by the ISI. These efforts must be done by co-operative effort from researchers from diverse countries.

ACKNOWLEDGEMENTS

The author thanks Adriano Baessa, Leandro Silva, Ana Luíza Lara, Ana Paula Verona, Camila Lins, Cíntia Oliveira, Fábio Salazar, Hérica Righi, Regina Fernandes and Túlio Cravo for research assistance. The collaborative work with Américo Bernardes in a previous paper is a precondition for this chapter. Support from FeSBE, CNPq and FAPEMIG is acknowledged. This chapter has benefited from comments, criticisms, and suggestions made by the editors of this book. The usual disclaimer holds.

REFERENCES

Aboites, J. (1996). *Analysis of patenting activity in Mexico* (Preliminary draft). Mexico (mimeo).

Albuquerque, E. (2000). Domestic patents and developing countries: arguments for their study and data from Brazil (1980–1995). *Research Policy*, 29 (9), 1047–1060.

Amsden, A.H. (2001). *The rise of "the rest": challenges to the West from late–industrializing economies*. Oxford: Oxford University.

Barro, R., Sala-I-Martin, X. (1995). *Economic growth*. New York: McGraw-Hill.

Bernardes, A., Albuquerque, E. (2003). Cross-over, thresholds and the interactions between science and technology: lessons for less developed countries. *Research Policy*, 32 (5), 867–887.

Coe, D., Helpman, E., Hoffmaister, A. (1995). *North–South R&D spillovers*. Cambridge, MA: NBER Working Paper 5048.

D'Costa, A.P. (2002). Uneven and combined development: understanding India's software exports. *World Development*, 31 (1), 211–226.

Dosi, G. (1988). Sources, procedures and microeconomic effects of innovation. *Journal of Economic Literature*, 27, Sept.

Fagerberg, J. (1994). Technology and international differences in growth rates. *Journal of Economic Literature*, 32, September.

Griliches, Z. (1990). Patent statistics as economic indicators: a survey. *Journal of Economic Literature*, 28, 1661–1707.

Grupp, H. (1998). *Foundations of the Economics of Innovation*: *theory, measurement and practice*. Cheltenham: Edward Elgar.

Lattimore, R., Revesz, J. (1996). *Australian science: performance from published papers*. Bureau of Industry Economics, Report 96/3, Canberra: Australian Government Printing Office.

MIT/SOFTEX (2002). *Indústria de software no Brasil*: *fortalecendo a economia do conhecimento*. Campinas: SOFTEX.

Observatoire des Sciences et des Téchniques (2000). *Science & Technologie*: *indicateurs 2000.* Paris: Economica.

OECD (2002). *Benchmarking industry–science relationships*. Paris: OECD.

Patel, P., Pavitt, K. (1995). *Patterns of technological activity: their measurement and interpretation*. In P. Stoneman (Ed.), Handbook of the Economics of Innovation and Technological Change. Oxford: Blackwell.

Pavitt, K. (1998). The social shape of the national science base. *Research Policy*, 27 (8), 793–805.

RAJESWARI, A.R. (1996). Indian patent statistics — An analysis. *Scientometrics*, 36 (1), 109–130.

Sandelin, B., Sarafoglou, N. (2003). Language and scientific publication statistics: a note. *Language problems and language planning* (forthcoming).

Silva, L. (2003). *Padrões de interação entre ciência e tecnologia.* Dissertação de Mestrado. Belo Horizonte: Cedeplar-UFMG.

The Government of the Republic of South Africa (2002). *South Africa's national research and development strategy*. Pretoria: The Government of the Republic of South Africa.

Velho, L. (1987). The author and the beholder: how paradigm commitments can influence the interpretation of research results. *Scientometrics*, 11 (1–2), 59–70.

Viotti, E., Macedo, M.M. (2003). *Indicadores de ciência, tecnologia e inovação no Brasil.* Campinas: Editora Unicamp.

World Bank (2003). World Bank Indicators (available at www.worldbank.org/wdi).

About the Authors

Birgitte Andersen is Senior Lecturer and Director for E-commerce Programme in the Department of Management at Birkbeck, University of London (UK). Her research and publication profile includes the economics, management and political economy of intellectual property rights (IPRs), as well as evolutionary economics and industrial dynamics with respect to innovation, institutions and the new economy. Previously she conducted her Ph.D. at the University of Reading, and she worked as Research Fellow at the ESRC Centre for Research on Innovation and Competition (CRIC), The University of Manchester. She has owned and/or worked on different international research projects funded by the Danish Research Academy, the British Academy, the Arts and Humanities Research Board, the Danish Social Science Research Council (SFF), the British Economic and Social Research Council (ESRC), and the European Union (EU). On issues related to IPRs and the productivity paradox, she has collaborated with economists from International Organisations as well as policy makers within and outside the European Union.

Subbiah Arunachalam is a volunteer and Distinguished Fellow at the M S Swaminathan Research Foundation, Chennai, India, since April 1996. He is an Honorary Fellow of the Chartered Institute of Library and Information Professionals and an Honorary Member of the American Society for Information Science and Technology. He is on the International Advisory Board of IICD, The Hague, chairman of the Board of Trustees of OneWorld South Asia and a Stockholm Challenge Award Ambassador. He is on the editorial boards of *Journal of Information Science*, *Scientometrics*, *Current Science*, *Current Contents* and *Electronic Journal on Information Systems in*

Developing Countries. His research interests include scientometrics, science in the developing world, and information technology enabled sustainable development.

Rémi Barré holds an Engineering diploma from Ecole des Mines (1971) and a Doctorate in Economics from Ecole des Hautes Etudes en Sciences Sociales (Paris, 1982). After having set up a consulting firm in economic analysis and modelling, he worked for the Ministry of Research in Paris where he developed strategic analysis and prospective activities. This led to the creation of the Observatoire des Sciences et des Techniques (OST), a public institution in charge of the production of S&T indicators, which he directed from 1990 to 2001; in parallel he worked with OECD and the European Commission on policy–related topics. Since 2002, he is Professor of Science Policy at the Conservatoire National des Arts-et-Métiers (CNAM), which is the University for continuing education in Paris. He is associated to the national foresight exercise FutuRIS.

Elise Bassecoulard, born 1945, dipl. Engineer and Master in Management Sciences, is a member of the IBIS scientometrics team of the INRA-Lereco Laboratory (Nantes, France). She is particularly involved in application of bibliometrics for INRA. She forms a team with **Michel Zitt**, born 1947, Ph.D. in Management Sciences and Docteur d'Etat in Economics. He leads the IBIS scientometrics team of the INRA-Lereco Lab and works as a consultant at the Observatoire des Sciences et des Techniques (OST, Paris). OST is in charge of S&T indicators for France in the European and international context. He has been more particularly involved in the development of science indicators. He is a member of the editorial board of the journal *Scientometrics* and of scientific committees of international conferences cycles in the field (ISSI, S&T Indicators). The team's general interests are science internationalisation mechanisms, international and inter-regional collaboration and spatial dimensions of S&T activities, long-term evolution of research systems, and science–technology relationships. On the methodological side, they worked on developments in citation analysis especially co-citation mapping, indicators at the journal level, interpretation and calibration of bibliometric series. They are currently working on publication–patents informetric relations, citation–based measures of excellence and, within the European network Prime, on dynamics of field emergence. They mainly published in *Scientometrics* and also in *Journal of the American Society for Information Science and Technology*, *Research Policy*, and *Science and Public Policy*.

Lennart Björneborn is Assistant Professor at the Royal School of Library and Information Science, Copenhagen, Denmark, since 2003. He obtained his Ph.D. in 2004. His Ph.D. Thesis is on *Small-World Link Structures across an Academic Web Space – a Library and Information Science Approach.* His research areas are: webometrics, link structure analysis, small-world connectivity, distributed knowledge organisation, serendipitous information finding, and networked creativity stimulation. He has published papers on webometrics with Peter Ingwersen in *Scientometrics* (2001) and *Journal of the American Society for Information Science and Technology* (forthcoming), and with Mike Thelwall and Liwen Vaughan in *Annual Review of Information Science and Technology* (forthcoming).

Andrea Bonaccorsi is Professor of Economics and Management at the School of Engineering of the University of Pisa (Italy). He has published several books and articles in *International Journal of Industrial Organization*; *Structural Change and Economic Dynamics; Research Policy*; *Research Evaluation*; *Scientometrics*; *Economics of Innovation and New Technology*, among others. He is a member of the Executive Committee of the PRIME Network of Excellence on science economics and policy (6th EU Framework Programme). He has served as advisor for the Italian government, for public institutions and for international organisations in problems of research and innovation policy.

María Bordons has a background in Medicine, in which she obtained a Ph.D. at the Universidad Complutense de Madrid (Spain). She joined CINDOC in 1994, as tenured scientist. Since then, she has worked in numerous research projects financed by different Spanish agencies. She has been involved in teaching at MSc and Ph.D. programmes, especially in the Universidad Carlos III de Madrid. Her academic papers have been published in journals such as *Scientometrics*, *Research Evaluation* and *Journal of the American Society for Information Science and Technology*. Her main research interests focus on the usefulness of bibliometric indicators for research evaluation, the study of collaboration, interdisciplinarity and research teams through quantitative approaches, and women and science issues.

Tibor Braun is Professor of Chemistry at the Lorand Eötvös University and Director of the Information Science and Scientometric Research Unit (ISSRU) of the Hungarian Academy of Sciences, Budapest, Hungary. His Ph.D. and DSc were defended in Budapest in 1967 and 1980. He has published more than 300 papers and 26 books, and has been awarded the international George Hevesy Award, the international Derek de Solla Price Award and the Chemistry and the Life Performance Awards of the

Hungarian Academy of Sciences. He is the founder of three international journals *(Journal of Radioanalytical and Nuclear Chemistry, Scientometrics,* and *Fullerene Science and Technology*), and is still the Editor–in–Chief of the first two of them.

Anthony Breitzman is Vice President and CTO of CHI Research, Inc. He has worked extensively with Fortune 500 companies helping them to understand and exploit their technological strengths. He is also responsible for designing and managing the technical direction of CHI. He has published widely in technology evaluation and mathematical techniques, and has presented at numerous conferences worldwide. He holds a BSc in Mathematics from Stockton State College; an MA in Mathematics from Temple University; and a Ph.D. in Computer Science from Drexel University.

Stefano Breschi obtained his Ph.D. from the Department of Economics, Università di Pavia (Italy) in 1994. He is Associate Professor of Industrial Economics and Deputy Director of CESPRI, Bocconi University, Milan, Italy. His main research interests are in the Economics of Technical Change, Industrial Organisation and Industrial Dynamics, Theory of the Firm and Networks, Economic Geography and Regional Economics. He participated in many national as well as international research projects of the European Commission and other institutions. Among his most recent publications: “Technological Regimes and Schumpeterian Patterns of Innovation”, *Economic Journal* 110 (463), 2000 (with F. Malerba and L. Orsenigo); “Networks of inventors and the location of academic research: An exploration of Italian data”, *Research Policy* (forthcoming, with M. Balconi and F. Lissoni); and “Knowledge spillovers and local innovation systems: A critical survey”, *Industrial and Corporate Change* 10 (4), 2001 (with F. Lissoni).

Stefano Brusoni is Assistant Professor (Research) at Bocconi University (Milan, Italy). He was awarded his Ph.D. by SPRU, University of Sussex in 2001. His research interests include the division and coordination of innovative labour; the organisational implications of knowledge specialisation and product design strategies; modularity; and the measurement of knowledge production and distribution processes within and across firms. He has published his work in *Administrative Science Quarterly*, *Research Policy*, *Industrial and Corporate Change*, *Journal of Management Studies, Economics of Innovation and New Technology*. He has extensive teaching experience in innovation management, strategy, organisation sciences, and business economics.

Linda Butler graduated in economics from the University of Queensland in 1972. She joined Paul Bourke at the Australian National University (ANU) in 1988, where they collaborated with Ben Martin and Jim Skea at the University of Sussex on developing bibliometric indicators at sub-institution levels. Since 2000 she has been head of the Research Evaluation and Policy Project at ANU. As well as continuing her research on evaluative bibliometrics, and research evaluation more generally, she has undertaken several consultancies for major research agencies, and her expertise is regularly sought by government departments and agencies whose responsibilities span research policy.

Rickard Danell is senior lecturer in sociology and information science at Umeå University (Sweden), were he received his Ph.D. in sociology in 2001. His research interests are bibliometrics and sociology of science. He is currently studying the development of Swedish science, with special focus on the changing labour market for Swedish Ph.D's.

Cinzia Daraio is post-doc research fellow at the IIT-CNR (Italy). She obtained a Master of Arts in Statistics at the Université Catholique de Louvain (UCL, Belgium), a *Diploma* in Economics and a Ph.D. in Economics and Management at the Scuola Superiore Sant'Anna (Pisa, Italy). She has research collaborations with the Scuola Superiore Sant'Anna and the Institute of Statistics, Université Catholique de Louvain. She is an active member of the PRIME working group on Public Sector Research and is involved in several international research projects. Some of her works have been published in international refereed journals like *Journal of Productivity Analysis*, *Research Evaluation*, and *Scientometrics*.

Koenraad Debackere is a professor of Technology and Innovation Management at the Faculty of Economics and Applied Economics at Katholieke Universiteit Leuven (Belgium). He is promotor of "Steunpunt O&O Statistieken", a research center funded by the Flemish government that is involved in studies on science, technology and innovation indicators. His research interests focus on the interface between innovation management and innovation policy.

Aldo Geuna is a Senior Lecturer at SPRU, University of Sussex (UK). He has been a Research Fellow at BETA, Université Louis Pasteur (Strasbourg 1), and at MERIT, Maastricht University. His research interests include: Economics of Knowledge Production and Distribution, Economics of Innovation and Technological Change, and Science and Technology

Policy. He has published in leading international journals in the area of university research and science–technology relationships. In 1999 he published *The Economics of Knowledge Production*, with Edward Elgar. He edited with Ammon Salter and Ed Steinmueller the book "*Science and Innovation: Rethinking the Rationales for Funding and Governance*" published by Edward Elgar in 2003.

Wolfgang Glänzel works as Senior Research Fellow at the Steunpunt O&O Statistieken at the Katholieke Universiteit Leuven (Belgium), and is also Senior Scientist at the Institute for Research Organisation of the Hungarian Academy of Sciences (Budapest). His main research interests are quantitative analyses and models of information processes in scientific research, and the theory of probability distributions. In 1999, he received the *Derek de Solla Price Award* for outstanding contributions to the quantitative studies of science. He is Co-editor of the journal *Scientometrics* and Secretary/Treasurer of the *International Society for Scientometrics and Informetrics (ISSI)*.

Isabel Gomez, Ph.D. in Sciences, is Research Professor of the Spanish Council for Scientific Research (CSIC). She started her research activities in experimental Biochemistry and later moved to CINDOC, where she works in Science Policy and Information Science. Her main field of specialisation is Bibliometric Indicators. She conducted research on the identification of research groups with bibliometric tools; interdisciplinarity in research, evaluation of national scientific journals; domestic and international collaboration networks in different fields of science; links between scientific and technological indicators; and analyses of the Spanish and Latin American scientific output for Research Management purposes.

Hariolf Grupp born 1950, Master (1975) and Doctor (1978) in Physics and Mathematics, Heidelberg University (Germany). Habilitation (professorial dissertation) in Economics, Technical University Berlin (1997). Senior researcher at the Scientific Office of the German Parliament (1980–1984). Since 1985 at the Fraunhofer Institute for Systems and Innovation Research (ISI), deputy director of this institute since 1996. Full Professor for System Dynamics and Innovation and Director of the Institute for Economic Policy Research (IWW), Economics Faculty, Karlsruhe Technical University (since 2001). Main fields of present research: Economics of technical change, science and technology indicators, industrial R&D management and research policy, and technology foresight.

Dominique Guellec is chief economist at the *European Patent Office*. He informs EPO policy for economic related matters and conducts studies on economic aspects of patenting in Europe. Until December 2003 he was senior economist at OECD. He was responsible for the science and technology indicators unit, which sets methodologies, harmonises and publishes S&T statistics from 30 industrialised countries, regarding notably R&D, human resources, innovation and patents. He has contributed to various OECD studies regarding technological change and economic growth, and published various books and academic articles on the economics of growth, innovation and patenting.

Diana Hicks is Professor and Chair of the School of Public Policy, Georgia Tech (USA). For almost 10 years she was on the faculty of SPRU, University of Sussex in the U.K. She was previously the Senior Policy Analyst at CHI Research, Inc. She has conducted quantitative assessments and served as a consultant for the Advanced Technology Program of the National Institute of Standards and Technology, the American Cancer Society, the Council for Chemical Research, the Department of Energy, the heads of the UK Research Councils, the Japanese National Institute for Science and Technology Policy, the National Science Foundation, the Small Business Administration, and The Royal Society in the UK. Prof. Hicks has taught at the Haas School of Business at the University of California, Berkeley and worked at the National Institute of Science and Technology Policy (NISTEP) in Tokyo. She is an honorary fellow of the Science Policy Research Unit, University of Sussex, UK.

Sybille Hinze was born in 1966. In 1990 she graduated in management of science from Humboldt University, Berlin. From 1990–1997 she was a research fellow at Fraunhofer ISI. In 1997 she completed her Ph.D. thesis at Leiden University at the Centre for Science and Technology Studies (the Netherlands). During 1997–1999 she was a postdoctoral fellow at the Australian National University, Canberra (Australia). Since 1999 she is a senior research fellow at Fraunhofer ISI engaging in quantitative studies, in particular in using and developing patent and bibliometric indicators. Her research activities are focussing on analysing systems of innovation in the life sciences.

Peter Ingwersen is since 2000 Research Professor at the Royal School of Library and Information Science, Copenhagen, Denmark. Ph.D. in 1991. Head of Department of IR Theory, 1993–2000. Visiting scholar at the European Space Agency, Italy, 1980–84. Visiting Professor at the Dept. of Information Studies, Tampere University, Finland, 1999–2002. Research areas: Interactive IR; Evaluation methods for work task–based IR; *Informetrics-Scientometrics & Webometrics*. He has published several books, and more than 60 journal articles and conference papers, in addition to editing work. Among his academic awards are the 1993 Jason Farradane, UK; the 1994 ASIS New Jersey Chapter, USA, Distinguished Lectureship Award and, in 2003, the Distinguished Research Award, American Society for Information Science & Technology.

Bihui Jin is senior research fellow at the Documentation and Information Centre of the Chinese Academy of Sciences, and professor at Nanjing University. She is the chief engineer of the Chinese Science Citation Database (CSCD) and leader of the scientometric research group at DICCAS. She is director of the international communication committee of the Chinese Association for Science of Science and S&T Policy. Bihui Jin has published several articles in international and national journals, and is responsible for drafting a number of Chinese government documents on S&T evaluation. She is a member of ISSI and was co-organiser of 9th International ISSI Conference on Scientometrics & Informetrics (Beijing).

Thed van Leeuwen, political scientist from the University of Amsterdam, started working as a researcher at the Centre for Science and Technology Studies (CWTS) at Leiden University in 1989. From the early 1990s, he became involved in research evaluation studies conducted by CWTS for different ‘actors’ in the Dutch science system. His research focused on the methodology of research assessments, and on the validity and applicability of journal impact measures. The study of the empirical data resulting from the bibliometric practice at CWTS on both research evaluation procedures as well as journal impact measures constituted in May 2004 the basis for his Ph.D. thesis.

Edda Leopold has earned Diploma degrees in mathematics and musicology in March 1989 and June 1990 from the University of Gießen (Germany). From January 1991 to December 1995 she was research Fellow at the Department of Computational Linguistics of the University of Trier. She received her Ph.D. in quantitative linguistics in July 1998 from the University of Trier. Since May 1999 she is senior researcher in the Knowledge Discovery Team of the Fraunhofer Institute for Autonomous Intelligent Systems in Sankt Augustin. She is currently working on the application of different data mining techniques to text and multimedia mining problems. Her research interests include text- and multimedia categorisation, spatial representation of meaning, multimedia semiotics, and multimodal communication.

Grant Lewison is Visiting Professor at City University (London, UK), and has been working in bibliometrics since 1988. He has a Ph.D. from the University of Cambridge in hydrodynamics, and worked on ship dynamics for nearly 20 years before switching to science policy in the UK Department of Trade and Industry and the European Commission. From 1993–2003, he was Senior Policy Adviser at the Wellcome Trust and developed a database of UK biomedical research papers with their funding acknowledgements on behalf of the Trust and a club of other research funders.

Francesco Lissoni is Associate Professor of Applied Economics at the University of Brescia, Italy, specialising in Economics of Innovation, and Deputy Director of CESPRI, Bocconi University (Milan, Italy). He graduated from Bocconi University in 1990, and undertook his graduate study at the School of Economic Studies of the University of Manchester, UK, 1992–95. His publications deal with the adoption and diffusion of innovations, the economics of technology transfer, and the geography of innovation. His work at CESPRI includes the creation of a number of patent–based datasets on technological activities by firms and individuals in the US, Europe and Japan.

Marc Luwel is a Ph.D. in physics. On leave from his position as senior researcher at the Flemish Ministry of Science and Innovation, he is an executive board member of the Dutch Flemish Accreditation Organisation and guest researcher at the Centre for Science and Technology Studies at the University of Leiden. He has been deputy chief of staff of the Flemish Minister of Education and advisor to the Belgian Prime Minister. Areas of interest and current work: quantitative studies to support the conception and evaluation of science and innovation policy, bibliometric analysis of research performance, knowledge flows between science and technology.

Daniela Luzi has obtained a degree in Modern Language and Literature with specialisation in Documentation and Librarianship. She is researcher of the Italian Research Council at the Institute of Research on Population and Social Politics. Her research interests have concentrated upon the analysis of the impact of Information and Communication Technologies on information management and communication processes among scientists. Within this context research topics such as Grey Literature, Open Archives, scientific evaluation methods and practice have been dealt with in scientific papers and publications. She taught Documentation at the University of Naples and at present she holds seminars on Information and Documentation Theory and Practice.

Michael May works at the Fraunhofer Institute for Autonomous Intelligent Systems (AIS, Germany), is head of the Knowledge Discovery Team, and leads the research efforts at the intersection of Data Mining and Knowledge Discovery, visual knowledge exploration, and databases. His current research interests are Spatial Data Mining; the integration of Data mining algorithms in databases and data warehouses; and the formalisation of causal reasoning processes. He is currently coordinator of a project integrating approaches from data mining and Geographic Information Systems in a Spatial Data Mining platform, and of an European Network of Excellence in Knowledge Discovery.

Henk F. Moed obtained a degree in Mathematics at the University of Amsterdam in 1978, and a Ph.D. in Science Studies at Leiden University in 1989. From 1981 he is a senior staff member at the Centre for Science and Technology Studies (CWTS) at Leiden University. In 1999 he was recipient of the Derek de Solla Price Award. He was Program Chair of the 5th (Hinxton) and 6th (Leiden) International Conference on Science and Technology Indicators. He published over 100 research articles, proceedings contributions and reports on topics in quantitative assessment of the research performance of research groups and institutions; research evaluation; the structure of national academic systems; the effectiveness of research funding; international scientific collaboration; and scientific journal performance and journal impact factors.

Mary Ellen Mogee is a career science and technology policy analyst with more than 30 years experience. She is currently president of Mogee Research & Analysis, LLC, a research and consulting firm specialising in patent analysis for technology assessment and competitive intelligence. Before founding her firm, Dr. Mogee was a technology policy analyst in the US federal government. She has spoken and written widely on the subjects of technology indicators and technology policy. She has taught science and technology management at the graduate level. She has a B.A. in Chemistry from MacMurray College; an M.A. in Science, Technology, and Public Policy from George Washington University; and a Ph.D. in Political Science also from George Washington University.

Fernanda Morillo obtained a Ph.D. in Information Management at the Universidad Carlos III de Madrid. She began her training with a scholarship at the Bibliometrics Department of CINDOC, where she actually works as a Technician. She has participated in research activities related to studies on interdisciplinarity of the sciences, and on the structure and dynamics of research groups in several scientific disciplines through the use of advanced bibliometric indicators. She is involved in the elaboration of bibliometric indicators of scientific production for Spain and Latin American countries, and data exploitation at different levels.

Eduardo da Motta e Albuquerque is Associate Professor of Economics at Universidade Federal de Minas Gerais (UFMG) in Belo Horizonte, Brazil. He teaches Economics of Science and Technology, Industrial Economics, and Development Economics, both in undergraduate and graduate courses. He is Doctor in Economics (Universidade Federal do Rio de Janeiro, 1998), having spent one year at SPRU–University of Sussex (TAGS Programme, 1996–1997). His papers have been published in leading Brazilian economic journals (*Revista de Economia Política*, *Revista Brasileira de Economia*, *Estudos Econômicos*, among others) and in *Research Policy*. His research subjects are national systems of innovation, catching up processes, and the role of science in development.

Fulvio Naldi obtained in 1976 a degree *cum laude* in Physics. He is presently Chief Engineer at the Italian Research Council, with the task of improving co-operation between public research and private industry. For more than 15 years he worked in the field of documentation of research projects and scientific output. Coordinator of several national and international projects for the implementation of information systems about the public funded research and member of three EC Committees for the exchange and the harmonisation of data on research activities. From 1999 to 2001, together with Ilaria Vannini Parenti, he has been project leader for the EC projects 'Patents and Bibliometric Indicators by Gender'.

Francis Narin is President of CHI Research Inc., which he founded in 1968. He is recognised as a leading expert on science and technology analysis, and he has acted as an expert witness in patent litigation. He is one of the pioneers of patent citation analysis, and developed many of the key techniques in this area. His work has been published widely, and he has presented his work at conferences around the world. He holds a BSc in Chemistry from Franklin and Marshall College; an MSc in Nuclear Engineering from North Carolina State College; and a Ph.D. from Walden University.

Lionel Nesta is Research Fellow at SPRU, University of Sussex. He is a Ph.D. student in Applied Economics with research interests in the Economics of Knowledge, Industrial Organisation, Economics of Innovation and Technological Change. His research has used and produced several indicators for the scientific and technological characteristics of organisations and nations based on large bibliometric databases.

Nils C. Newman's major research interests focus on the development of analytical tools for management of technology. He is the president and founder of IISC (www.iisco.com) and Director of New Business Development for Search Technology, Inc., Norcross, GA, USA. Mr. Newman has a Bachelor of Mechanical Engineering and an MS in Technology and Science Policy from Georgia Tech. He is currently pursuing a Ph.D. in Economics at the MERIT Institute of the University of Maastricht.

Ed Noyons has a degree in Dutch Linguistics at the de Vooys Institute of the University Utrecht, the Netherlands (1989). Since then he is working at the Centre for Science and Technology Studies (CWTS) of the Leiden University. In 1999 he received his Ph.D. degree at the Faculty of Social Sciences on the subject mapping of science in a science policy and research management context. As a research fellow at CWTS, he manages the science mapping projects and research on the development of an automated mapping process as well as the development of tools to define research areas.

Gerhard Paaß studied Mathematics and Computer Science and received a Ph.D. in Economics at the University of Bonn (Germany). Since the mid eighties he has designed statistical and knowledge–based algorithms and tools for extracting structure from data. Among others he has worked on probabilistic Bayes networks, neural networks, probabilistic reasoning, bootstrap methods, Bayesian Markov Chain Monte Carlo procedures, and Bayesian decision theory. He has led a number of projects aiming at knowledge discovery, text classification and web mining, and is head of the text and multimedia mining group in the Fraunhofer AIS. He is adjunct Prof. of Neurocomputing, at the Queensland University of Technology, Brisbane and on the editorial board of the *International Journal of Uncertainty, Fuzziness and Knowledge Based Systems*.

Parimal Patel is a Senior Research Fellow at SPRU, University of Sussex (UK). He is an economist with 20 years experience of working on technology policy issues. His main research interests are the measurement and analysis of technological activities at the firm, sectoral and national level. His most recent work has focussed on the following broad themes: linkages between science and technology, internationalisation of R&D, national systems of innovation and the nature and extent of technology diversification amongst large firms. He is the author of numerous articles and has presented papers at several international academic conferences.

Olle Persson is professor in Library and Information Science at Umeå University (Sweden) where he also obtained his Ph.D. in sociology in 1980. He was among the first in Europe to develop bibliometrics for indicators of scientific and technological activity, impact and collaboration as well as cognitive mapping of research fields. His work has been documented in a large number of journal articles. Among other things he has developed a software (Bibexcel) for bibliometric analysis. He is currently the leader of the Nordic Research School in Library and Information Science (NORSLIS).

Alan L. Porter's major concentration is technology intelligence, forecasting and assessment. He is Director of R&D for Search Technology, Inc., Norcross, GA, USA. He is also co-director of the Technology Policy and Assessment Center at Georgia Tech, where he is Professor Emeritus of Industrial & Systems Engineering and of Public Policy. His BSc is from Caltech and Ph.D. from UCLA.

Bruno van Pottelsberghe is professor at the Université Libre de Bruxelles (ULB, Belgium). As holder of the Solvay S.A. Chair of Innovation he teaches courses on the economics of innovation and intellectual property. He is Vice-president of the Solvay Business School and Director of its MBA Programmes. His research concerns the effectiveness of S&T policies, the internationalisation of R&D, innovation competencies and the factors fostering entrepreneurship. He has been consultant at the OECD (DSTI) and visiting researcher/professor at the Columbia Business School (NYC), at the Research Institute of the METI (Tokyo), and at the Institute of Innovation Research of Hitotsubashi University.

Anthony F.J. van Raan is Professor of Quantitative Studies of Science, Leiden University, the Netherlands, and Director of the Centre for Science and Technology Studies. Ph.D. in physics, Utrecht, 1973; post-doctoral fellow University of Bielefeld; 1985 'field switch' to science studies. Editor of the *Handbook of Quantitative Studies of Science and Technology* (Elsevier) and of the international journal *Research Evaluation*. Main research topics: advanced bibliometric methods, mapping of science and technology, and science as a 'self-organising' cognitive ecosystem. In 1995 he received, jointly with the American sociologist Robert K. Merton, the Derek de Solla Price Award.

.

Ronald Rousseau is professor at the KHBO, Department of Industrial Sciences and Technology, and guest professor at the Antwerp University (UA, Belgium), School for Library and Information Science. He has written numerous articles dealing with citation analysis, research evaluation, informetric models, information retrieval, library management, and web research. He is a member of the American Society for Information Science and Technology, and an active member of ISSI. Recipient of the Prize of the Royal Flemish Academy of Belgium for Science and the Arts (1979) and of the Derek de Solla Price award (2001). He is a honorary professor at Henan Normal University (P.R. China).

Bhaven Sampat received his doctorate in economics from Columbia University (USA) in 2001, and since then has been an assistant professor at Georgia Tech's School of Public Policy, where he won the "Faculty Member of the Year" award for academic years 2001–2 and 2002–3. From 2003–2005 he is on leave at the Department of Health Management and Policy at the University of Michigan. Dr. Sampat's research focuses on how scientific and technological changes affect the economy, how economic forces reciprocally affect the processes of scientific and technological change, and how to design public policies supporting and regulating science and technology.

Ulrich Schmoch studied mechanical engineering at the Universities of Hannover, Germany, and Grenoble, France, and graduated in 1977. He completed his Ph.D. thesis in the social sciences at the University of Hannover in 1993, and his habilitation thesis in sociology at the University of Karlsruhe in 2002. 1983 through 1985, he worked at a patent attorney's office in Wuppertal, Germany. In 1986 he entered the Fraunhofer Institute for Systems and Innovation Research in Karlsruhe, Germany, and currently he has the position of head of department. His research covers quantitative indicators, IPR policies as well as the interaction between universities and industrial enterprises.

András Schubert is Senior Scientist at the Institute for Research Organisation of the Hungarian Academy of Sciences (Budapest). His main research interest is quantitative modelling of social phenomena, most particularly the mechanism of scientific research.

He is Editor of the journal *Scientometrics*. He won the Derek de Solla Price Award for his outstanding contributions to the quantitative studies of science in 1993. He authored/co-authored more than 100 publications in scientometrics and related topics, and he is one of the most cited authors of this field.

Patrick Thomas is a Senior Analyst at CHI Research, Inc. He has worked extensively as a project manager in science and technology evaluation, and is a central figure in developing CHI's products marketed to the investment community. He has published widely in science and technology evaluation, and has presented at numerous international conferences. He holds a BSc in Management Science from UMIST; an MSc in Computer Science from the University of Birmingham; and a Ph.D. in Management Theory and Bibliometrics from Nottingham Trent University.

Robert J.W. Tijssen is Associate Professor at the Centre for Science and Technology Studies (CWTS), Leiden University, the Netherlands. Over the last 15 years he has been involved in numerous projects on S&T indicators for various government organisations, research councils, business enterprises, international research organisations, and for the European Commission. He has published numerous reports and research articles in international journals, and co-authored books on topics concerning to scientometric indicators and the evaluation of scientific performance. His current interests focus on empirical analyses of industry–university research linkages, inventors, and research excellence. He is the coordinator of the *Netherlands Observatory of Science and Technology* (www.nowt.nl), the producer of the bi-annual *Science and Technology Indicators Report* for the Dutch government.

Adriana Valente has obtained a degree *cum laude* in Law and degree *cum laude* in Social Communication. Researcher of the Institute of Research on Population and Social Politics of the Italian Research Council. Member of the Information Society Forum, Commission UE, from 1998 to 2001. Professor of Documentation and Scientific Communication at the Universities of Rome 'La Sapienza', Rome 'Lumsa', Naple 'Iun', Macerata, from 1989 to 2004. She has managed several projects and now she is coordinating the Italian Research Council–British Council project 'Perception and Awareness of Science'. She is author of more than 100 publications in the field of scientific communication.

Ilaria Vannini Parenti. Researcher at the Italian Research Council (CNR) Computer Centre in Milan, then responsible for the Documentation Service and Library, Research Area in Milan (CNR), and Director of the Publication and Science Information Office of CNR in Rome. Currently Consultant and project leader for a private Company. Milano Medicina award for scientometric studies of the scientific productivity of the Italian researchers compared with the international scene in the biomedical field. Project leader for the EC projects 'Patents and Bibliometric Indicators by Gender' together with Fulvio Naldi.

Arvids Ziedonis' broad interests are technology strategy, the economics of innovation, and technology policy. His current research is centered on understanding the factors that influence the acquisition and commercialisation of new technologies by firms, how institutional factors influence innovation, and the empirical measurement of innovative activity. Dr. Ziedonis received his Ph.D. from the University of California, Berkeley (USA) and is an assistant professor of corporate strategy and international business at the University of Michigan.

Subject Index