
Patent Data in Economic Analysis

Rafał Wisła

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/68100>

Abstract

The issues discussed in this chapter constitute a voice in a methodological discussion on the scope and manners of the utilisation of patent statistics in economic research. The discussion comprises the following issues: the gist of a patent monopoly, the evolution of opinions on the benefits and costs of a patent monopoly, and the possibilities and limitations of utilising patent statistics in the quantification of economic processes. This chapter is of a review and has methodological character. The analysis conducted within the text leads to two groups of conclusions. One of them concerns the shortcomings and limitations of patent databases, while the other concerns the identification of scientific exploration fields by means of patent metadata.¹

Keywords: patent data, patent information, patent statistics, patent statistics as economic indicators

1. Introduction

A patent is an exclusive right to use a new solution of a technical nature; it is considered as one of the strongest rights of intellectual property. In the scientific sense, a patent is the crowning point of research and development activities. In the economic dimension, it is one of the stages of the innovation process. From the point of view of the patent owner, it constitutes a resource and a potential market value. A patent has a relatively high capability for transformation into a production factor. The properties of a patent's description and the exclusive right itself (a patent understood in a narrow sense) cause a situation in which patent

¹The issues discussed in this chapter were taken also in R. Wisła (2014). The regional patterns of technical knowledge accumulation in Central and Eastern Europe countries, authored monograph, which was published by Polish Scientific Publishers PWN, Warsaw. The book is available in Polish only.

information constitutes a bridge between the results of the research and development (R&D) processes and their potential economic utilisation.

A patent is a body of accumulated scientific, technical, and industrial knowledge with potential to influence the course of economic processes. It is an economic category ascertained in both normative economics and positive economics. In the former case, it is considered on the plane of institutional solutions (an optimum patent policy, the effectiveness of patent systems, and external effects); in the latter, it is regarded as a measure of the dynamics and direction of changes in the economy. An important advantage of the time series of patent applications (and patent awards) is the possibility of their simultaneous use in at least four dimensions, that is time, space, an economic sector, and value.

Before the middle of the 1970s, the average annual number of patent applications (regardless of the mode of applications) had remained at a stable level. In the years 1975–2008, the average annual increase in the number of patent applications was 3.2%, while in the years 1995–2008, this rate rose to 4.9% [1]. If the latter period is extended until 2016, the annual growth exceeds the level of 5%.

The main factors stimulating this trend include (1) the replicability of a patent protection application concerning technical solutions within a single invention, (2) an increase in the effectiveness of research and development activities caused by pressure on the applicability of research results, (3) the emergence of new areas of technological development and/or the greater intensity of the utilisation of the existing ones, and (4) a heightened awareness of the importance of the formal protection of intellectual capital.

Consequently, there appear huge collections of structured data and information (databases of facts). In combination with the rapid technological development in the field of the IT infrastructure of data repositories and the new methods and techniques of data mining, they open up new opportunities for: (1) discovering previously unknown relationships and connections among data, (2) projecting the course of various processes, including economic ones, (3) determining the regularities of such processes, and last, but not least, (4) attempting to formulate general rules for their course, depending on conditions shaping the environment.

An important advantage of patents and collections of information on patents (databases) is their long-term availability (counted even in tens of years). The content of patent databases and long time series describing them allow the aggregation of data at any (microeconomic, mesoeconomic, macroeconomic, or international) level. Patent databases can be used in different ways and for different purposes.

The main reasons for using patent databases include (1) growing demand for analytical work for the needs of science and technology policies, (2) acquiring industrial knowledge described in the patent literature, (3) monitoring patent activities (input resources for future innovative activity), (4) searching for and identifying the directions and dynamics of development trends in particular areas of technology, (5) evaluating the results of scientific and industrial research, and (6) mapping research and development centres (as well as other entities) with respect to cooperation and identifying cooperation networks.

The new possibilities and methods of creating, collecting, using, transmitting, and processing data, information, or knowledge cause an exponential increase in the supply of their resources. Their acquisition frequently takes place through multifunctional repositories combined with a modern system of services ensuring the acquisition of and access to their resources. From the point of view of social development and the increasing competitiveness of science and economy, such repositories constitute a powerful accelerator for the growing intensity and effectiveness of scientific research. Through access to various facilities, frequently extensive collections of sources, and the integration of distributed databases, they facilitate access to and productive utilisation of their resources.

The abundance of patent descriptions and patent statistics is not utilised sufficiently in the cognitive process in research into economic mechanisms and phenomena.

The topic of patent statistics and its use in economic research is not raised too frequently in the academic literature around the world. The intellectual foundations for the usefulness and possibility of using patent information collections in scientific research comprise the works of such researchers as: Pavitt [2, 3], Griliches [4], Jaffe, Trajtenberg, Henderson [5, 6], Schmoch [7–10], Guellec, van Pottelsberghe [11–13], Cohen, Merrill [14], Hall, Jaffe, Trajtenberg [15, 16], as well as OECD manuals [17], which harmonise the rules of patent statistics as one element of the system measuring technological changes, scientific and innovative activity, as well as the structural changes of the economic environment.

This chapter is organised as follows: Section 2 contains an explanation of the essence of the patent monopoly and presents opinions on its advantages and costs. Section 3 presents the main sources of patent information, the mode of making patent information generally available, the potential methods of using patent metadata, as well as fields of scientific exploration, using this category of source data. Section 4 includes scholarly reflection, based on the evolutionary approach, on the possibility of using patent statistics in economic research. Section 5 comprises a discussion on the methodological conditions for the utilisation of patent information. The issues presented in this chapter constitute participation in a methodological discussion on the scope and manners of the utilisation of patent statistics in economic research.

2. Patent

A patent is a personal property right which is effective towards all, transferable, and inheritable. It constitutes “property” within the meaning of the civil law. The making, using, offering, and marketing of a product and/or the manner constituting the subject matter of an invention are activities covered by exclusivity (the patent monopoly) resulting from the essence of a patent. The material scope of a patent is determined by patent claims included in a patent description (a description of an invention, drawings, constitutional formulae, structures of relationships, sequences, etc.). A patent document contains descriptions of a protected solution together with patent claims, constituting (synthetically formulated) scope of granted protection as the result of comparing the current state of the art with the protected solution.

As an economic mechanism, the patent has been present in the scholarly discourse since the beginning of the development of the economic sciences. However, in the eighteenth century and the first half of the nineteenth century, opinions on the patent were expressed on the margins of the main disputes in the field of political economics: Smith [18], Say [19], de Sismondi [20], Lotz [21], Jakob [22], and Mill [23]. The subsequent two or three decades of the nineteenth century witnessed the vibrant development of the economic literature devoted to exclusive property rights and the presentation of arguments both for and against the patent monopoly. The discussion focused mainly on the following four constructs: the natural law, justification for a temporary monopoly, stimuli for further creative activities, and reward for making knowledge publicly available [24].

From the aforementioned discussion, one could draw not only a tentative but also a very general conclusion that a temporary patent monopoly should be permitted. This opinion was advocated, to a greater or lesser degree, by A. Smith, J. Betham, J.S. Mill, J.H.G. Justi, L.H. Jakob, and J.F.E. Lotz. The clearly opposing opinion was held, among others, by Simonde de Sismondi.

It is difficult to determine unambiguously whether the views of the economists of that period constituted an important reason for work on an international convention on the protection of industrial property initiated in Paris, France, in 1873. (It was mainly the industrial, political, and legal circles that were the most interested in the development and international unification of the patent law.) But this fact had a considerable impact on the dynamics of research on the patent monopoly to be conducted by economists in the subsequent decades. Nevertheless, for the record, it should be emphasised that such scholars as Fisher [25], Marshal [26], Vaughan [27], Clark [28], Plant [29], Robbins [30], Hayek [31], Nordhaus [32], Scherer [33], Horstmann, Macdonald, Slivinski [34], Baumol [35], Gilbert, Shapiro [36], Klemperer [37], Cohen, Nelson, Walsh [38], and Stiglitz [39] presented their positions on the architecture of the patent system (including the issues of exclusivity, territoriality, time limit). In the twentieth century, the system developed very quickly, generating a number of external effects (together with clearly intensifying negative effects).

Summarising, one could state that a patent fulfils the following two main functions: (1) protection, which is related to the controversial institution of the legal monopoly and (2) dissemination of knowledge, thanks to (structured) collections of the patent literature. For this reason, a patent (patent description) may be also understood as a scientific and technical publication similar to an article in an academic journal.

3. Patent databases

The main source of patent information is publicly available patent documentation (application descriptions, patent descriptions of inventions) which contains, first of all, information on the current state of the art in a given field. An important advantage of patent documents is their up-to-date character (in the worldwide sense) and the unambiguous legal status with respect to industrial property protection. Patent literature collections comprise

official bulletins of national offices and international organisations, bibliographic collections (metadata), as well as articles presenting particular problems, discussions, and past judicial decisions.

Patent information is provided under various procedures, but in practice, access to the full collection of metadata is not easy. The author has identified the following selected barriers against access to complete collections of patent data:

- (1) national patent offices do not provide functionalities and tools allowing one to acquire metadata automatically and in bulk,
- (2) public digital repositories of collected patent documentation have a relatively simple and functionally limited architecture,
- (3) reports drawn up by national and regional patent organisations and delivered to statistical offices are general and superficial; their subsequent visibility in public statistics (statistical offices) does not allow any serious research, and
- (4) commercial distributors try to overcome the aforementioned limitations; their acquisition of patent information is not only professional and functional but also expensive to the end user.

The general advantage of the time series (records) of patent applications (and patent awards) is the possibility of their utilisation in research on the development of science, technology, innovative activity, and structural changes in the economy [40–42] in at least four dimensions simultaneously: time, space, an economic sector, and the institution of property.

Rapid technological development in the area of the IT infrastructure of data repositories,² including patent information collections, is a strong factor accelerating increase in the quality, intensity, and effectiveness of scientific research.³ An important advantage of patent information collections is their long-term availability (counted even in tens of years). It offers ample opportunities for their utilisation in scientific research. The content of patent databases and long time series describing them allow the aggregation of data at any level. In the case of research on innovation conducted at the microeconomic, mesoeconomic, and macroeconomic levels, patent databases allow one to describe the following features of innovative activity:

- 1) the novelty level of products resulting from conducted research and development activities,
- 2) the types of innovations under development and technological competences,

²There exist two basic models of providing access to digital objects (records) in IT systems. One of them is remote access in which, during a harvesting process, metadata of resources remaining in the provider's repository are entered into the system repository and such metadata may be made available to the user and in the other model, material is placed directly in the system's repository base.

³The first researchers to discover these potential possibilities and to determine the direction of further research were Scherer (1965) and Schmookler (1966). Following the appearance of new technological possibilities (electronic data collections), Griliches (1984, 1990), Griliches, Pakes, and Hall (1987) started the empirical verification of their usefulness. Schankerman and Pakes (1986) were the first to work with data coming from European countries.

- 3) the sources of innovations, and
- 4) the dissemination of knowledge and technology.

Patent applications have been the subject matter of research processes for many years [4, 43–46]. What is frequently emphasised is the relationship among R&D activity, patents, and their impact on the stimulation of further R&D initiatives. Not all patent applications lead to the award of a patent. The difference between the number of applications and the number of patent awards may be used as a measure of the effectiveness of R&D activity.

Every patent provides a detailed description of an invention and is categorised into a particular class, group, or subgroup of the international patent classification. The hierarchical arrangement of the system facilitates research on patent applications with respect to novelty and inventiveness; it also allows precise research into technological trends at both the micro-economic level (innovations under development in a given corporation) and the macroeconomic level (the identification of the economy's technological advantages).

The dissemination of knowledge and technology may take place in the form of patents, unpatented inventions, licences, available know how, trademarks, projects, and designs. Attempts to measure the diffusion of knowledge and technology by means of patent databases or market transactions, or to identify relationships between producers of technical innovations and their users, have been made for at least 30 years. The relevant measurement methodologies developed so far emphasise various aspects of the discussion process, while the process of improving the quality of measuring the force of the dissemination of knowledge and technology is still far from its completion.

Thus, patent databases may be used in various ways. The number of patents awarded to a particular company, industrial sector, branch of the economy, region, and/or state reflects the level of technological dynamics. Examining the pace of changes, searching for relationships with particular patent classes or groups may help identify the directions and dynamics of technological changes.

Information included in patent information collections may be divided into the following three major pillars:

- 1) the technical specification of potential value of a new solution (a technical classification, the number of citations in other patent descriptions, the number of licences granted, the frequency of changes concerning the patentee),
- 2) the development of an invention (the structure of the team of inventors and their affiliations, the structure and types of applicants, progress in the development of a “triadic patent family”), and
- 3) the history of an application, which includes the application submission date (in a particular country, under other modes of the safeguard proceedings, etc.), the date of publication, the date of rejection or withdrawal, the date of patent award, the date of the expiry of the monopoly (failure to pay a fee or to extend a patent), etc.

The main international patent databases maintained and operated by various international organisations include the following:

- 1) *The European Patent Register, European Patent Register* and *Espacenet*: Databases maintained by the European Patent Office,
- 2) *Patentscope*: A database maintained by the World Intellectual Property Organisation,
- 3) *DEPATISnet*: A database and information service maintained by the German patent system,
- 4) *USPTO*: A full-text documentation base of patent applications and patents awarded in the United States of America, and
- 5) *Thomson Innovation*: A commercial database allowing the exploration of extensive and organised collections of patent applications and patent awards.

Other databases, usually with a particular thematic profile, include Cippix® (chemistry), GenomeQuest (biology), LexisNexis, MicroPatent and Delphion (integration of the USPTO, EPO, and WIPO databases), JP-NETe, and KPA Search In KIPRIS—Free Services.

The basic source of patent information is patent documents, that is published descriptions of inventions included in databases, which are open to the public. Patent databases include information on millions of submitted inventions. They constitute the richest source of information on the current state of technology in a particular field. Patent documents are (usually) published earlier than the technical literature.

At present, the most frequently used fields of patent statistics, remaining at the beginning stage of development, include the following:

- 1) patent family statistics, that is a stream of applications or a cluster of rights to an invention granted in more than one patent office; this phenomenon is characteristic of and intensifying mainly among entities operating in economically developed countries,
- 2) an evaluation of the monopoly's title, that is the value of a patent (different from the value of an invention itself) is estimated based on the scope of a patent family (the number and importance of patent offices, the geographical scope of patent protection), the number of citations in other patent descriptions, the length of the monopoly (incurring fees), the fact of initiating an objection/opposition procedure, the number and types of licences granted under a given exclusive right, the location of a protected solution in technology (significance for technological development⁴), the fact of establishing a new company whose business model is based on the patent monopoly,

⁴A company may have at its disposal a few alternative organisational techniques and processes related to product manufacturing, which may differ from one another in small details concerning the organisation and engagement of the production factors. A set of all available organisational techniques and processes related to the manufacture of final products is referred to as production technology. It is a technology specific for a particular company. A collection of all specific technologies constitutes the technology of a branch. The expansion of the collection of technologies in a particular company expands the technology of a given branch (Gomułka, 1998, pp. 12-13.)

- 3) patent statistics with respect to regions, and
- 4) patent statistics with respect to gender.

The weaknesses of patents as characteristic features of innovations are generally known. Many new or improved solutions are not submitted for patenting, while others are protected with numerous patents and/or other forms of protection. Many patents have no technological or economic value, while others are extremely valuable in this respect.

4. Patent statistics as an economic indicator

In 1990, the *Journal of Economic Literature* published an article entitled *Patent Statistics as Economic Indicators: A Survey* by Z. Griliches,⁵ who regarded technological changes as the main source of long-term economic growth. The narration and arguments used in this article can be characterised in one sentence as follows: “In this desert of data (necessary to describe the sources of economic growth, technological or structural changes, competitive positions, *author’s note*), patent statistics loom up as a mirage of wonderful plentitude and objectivity (i.e. qualities required of the time series of economic variables, *author’s note*)”. Similar studies on the possibilities of quantifying the relationships between technological changes and economic effects had been undertaken earlier by Schmookler [47, 48], Pavitt [3], Basberg [49], and Schankerman [50].

At the beginning of the 1950s, Schmookler [47] referred to a patent as a result of innovative activity. He identified the course of a patent activity trend (determined on the basis of the number of patent applications and patent awards) with some kind of an innovative activity indicator. In patent data collections, he searched for an explanation for the rising productivity of the American economy. However, what should be stressed is Schmookler’s considerable carefulness in this respect. In reality, it was difficult to observe any strong and repeatable relationship between the combined productivity of the production factors and the dynamics of patent activity. Therefore, Schmookler indicated the directions of the potential utilisation of patent statistics rather than a measuring methodology itself. However, it should be remembered that in the 1950s, there was no systematic collection of data on R&D expenditures; what was collected was selected (and dispersed) data on the employment of scientists and researchers as well as the movement of the highly qualified research personnel. Patent statistics remained practically the only database which could be used to describe technological or structural changes as well as competitive positions at the microeconomic and macroeconomic levels.⁶

⁵Zvi Griliches, 1930-1999.

⁶In 1963, the first conference of the science ministers of the countries belonging to the OECD was held. It coincided with the publication of the first methodological guidelines for the collection, processing, and presentation of data related to R&D—*The Frascati Manual*. In 1966, the British “Science Policy Research Unit” initiated its activities. This was the beginning of the multidirectional development of statistics in the area of science, technology, and innovation (S-T-I). Analysing the evolution of the S-T-I methodological approaches, one can easily notice functional changes in this category of public statistics. In the 1990s, statistics concerning science, technology, and innovation entered the period of rapid changes.

Despite these barriers, the early 1960s witnessed the beginning, and the subsequent decades the continuation, of the research programme which, from today's perspective, could be called "an analysis of the rate of return from investments in R&D". The researchers who were especially prominent in these first two decades were Zvi Griliches, Edwin Mansfield, Jacob Schmookler, and Nestor E. Terleckyj.

In the first half of the 1980s, Pakes and Griliches [51, 52] put forward an interesting theoretical construct whose aim was to explain the impact of knowledge created in the industry on the productivity of the production factors. In the analysed contexts, they classified knowledge as "technical knowledge of particular economic value (K), accumulated in a particular period of time $\dot{k} \frac{dk}{dt}$ ". In their original model, the explanatory variables of the category \dot{k} (of both the input character and output character), they pointed at: (1) expenditures on R&D activity, (2) expenditures on traditional capital goods, (3) patent activity, (4) the productivity of the traditional production factors, (5) the (market) value of a business enterprise.

In this model, a patent (patent activity) is an imperfect quantitative characteristic of a company's innovative activity in a very close relationship with \dot{k} , (technological accumulation, technological learning):

$$p_{i,t} = dt + \beta \dot{k}_{i,t} + v_{i,t}^* \quad (1)$$

where $p_{i,t}$ is a logarithm of quantitatively described patent activity, dt is a derivative of the function of the time trend, $v_{i,t}^*$ is an error uncorrelated with \dot{k} and with t , and β is the flexibility of patent activity with respect to \dot{k} (industrial knowledge accumulation, its direction, and dynamics).

Eq. (1) may be interpreted as a simplified model of patent activity.

The 1980s brought considerably greater opportunities for the empirical verification of associations between patent activity and other economic characteristics. Hausman, Hall, and Griliches [53, 54] looked for a standard relationship between expenditures on R&D and patent applications. They formulated four basic research questions concerning the following areas: (1) the strength of the relationship between R&D expenditures and patent applications (patent awards), (2) the flexibility of patent applications in response to changes in R&D expenditures, (3) the distribution of the time series of the effects of R&D activity, and (4) the singularity of the standards in the time courses of these relationships.

They verified their original econometric model empirically, using a sample of 128 business enterprises (1984); two years later, it was already a sample of 642 entities. The results of their research confirmed the hypothesis about the relationship between the examined companies' R&D expenditures and their patent activity.

However, depending on the size of a given company, its patent policy, and the previous results of R&D activity in correspondence with the effectiveness of conducted business operations, this relationship can have different levels of dynamics and strength.

At the same time, Pakes [55] studied relationships among companies' R&D expenditures, patent awards, and market valuation. The conducted research revealed that unexpected (for

the capital market) changes in R&D expenditures and patent activity caused considerable changes in the market valuation of companies.

Griliches [4] asked two fundamental and still relevant questions concerning the possibilities of using patents as an economic indicator. Firstly, which aspects of economic activity are in fact described by patent statistics? Secondly, what is to be measured by means of patent statistics? Despite such questions and justified doubts, he accepts the assumption that patent activity was a good indicator of the effectiveness of research and development activity.

He regarded R&D expenditures as a measure of contribution to inventive activity, while patents—as the result of this activity. He formulated a hypothesis on a strong relationship between R&D expenditures and the number of patent applications. In order to verify the hypothesis, he built the following knowledge production model [4]:

$$P = \alpha K + v = \alpha R + \alpha u + v, \quad (2)$$

where P is the patents as a quantitative measure of inventiveness or production of industrial knowledge, K is an unobservable variable expressing the net increase of economically valuable knowledge, R is expenditures on research and development invested in inventive activity, U is other sources of knowledge increase, v is a random component, and a is a structural parameter of the model.

According to the original concept, Griliches considered the parameter a standing next to K , R , and u as the same because he was forced to quantify K as follows: $K = R + u$; he had to look at the dynamics of the net increase of the economically valuable knowledge on the side of expenditures. Griliches's model was verified empirically. The main conclusion resulting from research on industrial knowledge production in the United States of America concerns the positive relationship, observable in the long period of time,⁷ between expenditures on R&D activity and the intensity of patent applications ($a = 0.76$).

However, nowadays, there is considerable space for the evaluation of the results of inventive activity and the evaluation of the force of their influence on the scientific, technological, and economic environments.

The rapid development of the IT infrastructure of patent databases which has continued since that period allows a relatively objective quantification of the value of a particular technical solution included in a patent description. The citation intensity of a particular patent description, information on granted licences, information on changes concerning the patentee, and the intensity of “triadic patent families” of the entity submitting an application are the main variables which are subjectable to such quantification. Hence, in its essence, k in Eq. (1) becomes more and more quantifiable.

The traditional evaluative approach to the economic quality and usefulness of industrial knowledge embodied in a new technical solution is a method based on the extension of the patent monopoly. Fees for subsequent periods of protection need to be paid in advance; a typical increase in fees for subsequent periods is described the best by means of the exponential function. It can

⁷The years were 1953–1989.

be assumed that for a typical business situation, maintaining the patent monopoly is economically justifiable. The longer the patent monopoly is maintained, the stronger (theoretically) the protected solution incorporates economic value.

Hall et al. [16] put forward an approach including the market valuation of a patent portfolio in correspondence to their citation intensity. They draw clear conclusions which, in fact, are the reflection of many years of their research in which patent information was used—(1) the number of citations of a particular patent claim in other patent claims is more important than an increase in the number of patent applications or patent awards, (2) the number of citations of a patent claim in other patent claims influences the market valuation of the patent holder (valuation of listed securities), and (3) the number of citations of a particular patent is a quantifiable manifestation of the diffusion of industrial knowledge.

5. Comments on the methodology of using patent statistics

A patent application is an economic event, one of the many stages in the innovation development process, (frequently) the crowning point of research and development activities. The acquired protective right constitutes a potential resource for an organisation's commercial activity which may evolve into a production factor. A patent is not an innovation, but its intermediate character causes a situation in which patent information constitutes some kind of a bridge between the results of a particular company's R&D activity and implementation activity.

The methodological discussion on the scope and methods of using patent statistics in economic research is not as intensive as the methodological discussions dedicated to innovation or bibliometry. Nevertheless, what results from the discussion is a catalogue of a few fundamental principles of designing research procedures.

Firstly, depending on a particular branch or sector of the economy, business enterprises are characterised by various expectations and strategies concerning the formal protection of industrial property. For example, industrial sectors with long cycles in which final products are created manifest particular prudence in ensuring long-term and strong systemic protection, while other sectors whose curve of demand and technology changes its position relatively quickly do not use the patent monopoly intensively. In this case, the business model is based mainly on the priority of rent. However, a comparative analysis of the same branches and/or sectors of the economy in a properly selected group of countries or regions is entirely justified. While macroeconomic and mesoeconomic analyses will be proper research tools, this will not be the case with juxtaposing business entities of different sizes operating in the same branch or sector. (This is so because the possibilities of acquiring and maintaining the patent monopoly by large business entities are very much different from those available to small and medium enterprises.)

Secondly, one should remember about the differences in patent procedures characteristic for particular cultures or legal systems. They constitute an important qualitative factor influencing

the number of patent applications and eventually awarded patent rights.⁸ This problem does not occur in countries following the procedure of a single regional application (e.g. the patent monopoly award procedure based on European patent applications).⁹

Thirdly, it needs to be emphasised that a considerable percentage of both patent applications and patent awards does not translate into any factual increase in the productivity of production factors. Hence, thorough analyses need to take into consideration such criteria as the number of citations of a particular patent description in other patent descriptions, information on granted licences, and information on changes concerning the patentee. The first criterion allows one to determine the value of a particular patent, while the other two, the factual utilisation of new solutions in production processes.

So far, the fulfilment of the condition of “thorough analysis” has been possible only in the case of data collected, processed, and made available by commercial providers of scientific and technological information in such countries as the United States of America, Germany, the UK, or France. The decisive majority of national patent systems do not collect information of this category or the content of patent descriptions is not attractive enough to be cited, and technical solutions themselves are not interesting enough for their property rights to be licensed.

Fourthly, analyses of the efficiency of national or regional patent systems or the effectiveness of entities submitting patent applications need to take into account time shifts. It is methodologically incorrect to conduct an evaluation within the range of the same year (patent applications vs. patent awards). Also, a simple juxtaposition of the time series of patent applications and patent awards does not result in any analytical content.

Fifthly, taking into consideration technological and economic criteria, patent awards are more valuable than patent applications. Therefore, the results on research on the distribution of the former are more relevant.

Sixthly, many stages of processing patent descriptions within a long patent procedure generate risks of the following types of errors: spelling errors (errors in the spelling of business names of applicants, names of inventors, etc.), factual errors (changes concerning formats and conventions of applicant registration, codes of the International Patent Classification, postal codes resulting from changes in the administrative division, etc.), and delays (registration of changes concerning patentees, granted licences).

6. Conclusions

The recent years have witnessed serious reflection on the factual and potential opportunities related to the use of patent statistics in measuring other processes. The appearance of

⁸For example, the “explosion” of patent applications under the national procedure in China in the years 2011–2016.

⁹Among organisations granting regional patent rights, one could mention the European Patent Organization, Eurasian Patent Organization, African Intellectual Property Organization, and African Regional Industrial Property Organisation.

such paradigms as the information economy, the knowledge-based economy, or the creative economy has been unavoidable. In these aspects, patent information becomes a relatively good reflection of the aforementioned development structures.

In 2010, the Patent Office of the People's Republic of China registered a 25% increase in the number of patent applications over that of 2009 [56]. This is just one of the many examples of the worldwide tendency in the field of patent protection. Consequently, there appear huge collections of data and information. Almost all such publicly accessible collections are adjusted mainly to the requirements of patent clearance analyses. It is frequently impossible to conduct any quantitative analyses based on such databases.

Attempts to bridge this gap are made by commercial providers of databases and analytical applications. Nevertheless, it is necessary to indicate basic information shortcomings characteristic for public registers which are used by commercial providers. Their elimination could improve significantly the quality of research based on the use of patent information. This includes collecting the following types of information: (1) the economic classification codes of entities submitting applications for patent protection (this will improve the effectiveness of sector-based research) and (2) notes on granted licences and changes concerning the patentee (this will allow research into the industrial property's secondary trading market; such information could be collected together with fees for the extension of the patent monopoly for the subsequent period).

In the author's opinion, there is still considerable space for scientific exploration based on the use of patent metadata. There are not only considerable risks but also potential benefits related to research based on patent metadata into the following topics: (1) the identification of the strength of synergy in the case of mergers and acquisitions, (2) cooperation networks and the diffusion of knowledge among business entities from different sectors, (3) the degree of globalisation of business activities, research teams, the development and dynamics of team structures, and the spatial mobility of scientific and industrial inventors, (4) economic forecasting, (5) the diffusion of technologies (based on the use of licence information included in databases), and (6) the industrial property secondary trading market.

Author details

Rafał Wisła

Address all correspondence to: rafal.wisla@uj.edu.pl

Jagiellonian University in Kraków, Institute of Economics, Finance and Management, Poland

References

- [1] World Intellectual Property Indicators (2011), WIPO Economics & Statistics Series. http://www.wipo.int/edocs/pubdocs/en/intproperty/941/wipo_pub_941_2011.pdf

- [2] Pavitt K. (1984), Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory, "Research Policy", 13 (6), pp. 343-373.
- [3] Pavitt K. (1988), Uses and Abuses of Patent Statistics. In A.F.J. van Raan, editor. Handbook of Quantitative Studies of Science and Technology. Amsterdam: Elsevier Science Publishers, pp. 509-535.
- [4] Griliches Z. (1990), Patent Statistics as Economic Indicators: A Survey, "Journal of Economic Literature", 28, pp. 1661-1707.
- [5] Jaffe A.B., Trajtenberg M., Henderson R. (1993), Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations, "Quarterly Journal of Economics", 108, pp. 577-598.
- [6] Jaffe A.B., Trajtenberg M. (1996), Flows of Knowledge from Universities and Federal Labs: Modeling the Flow of Patent Citations Over Time and Across Institutional and Geographic Boundaries, "Proceedings of the National Academy of Sciences", 93, pp. 12671-12677.
- [7] Schmoch U. (1993), Tracing the Knowledge Transfer from Science to Technology as Reflected in Patent Indicators, "Scientometrics", 26 (1), pp. 193-211.
- [8] Schmoch U. (1997), Indicators and the Relations between Science and Technology, "Scientometrics", 38 (1), pp. 103-116.
- [9] Schmoch U. (1999), Impact of International Patent Applications on Patent Indicators, "Research Evaluation", 8 (2), pp. 119-131.
- [10] Schmoch U. (2008), Concept of a Technology Classification for Country Comparisons. Final Report to the World Intellectual Property Organisation, Fraunhofer Institute for Systems and Innovation Research, Karlsruhe.
- [11] Guellec D., van Pottelsberghe B. (2000), Applications, Grants and the Value of Patents, "Economic Letters", 69 (1), pp. 109-114.
- [12] Guellec D., van Pottelsberghe B. (2001), The Internationalisation of Technology Analysed with Patent Data, "Research Policy", 30 (8), pp. 1256-1266.
- [13] Guellec D., van Pottelsberghe B. (2007), The Economics of the European Patent System: IP policy for innovation and competition. Oxford, New York: Oxford University Press. <http://lib.ugent.be/catalog/rug01:001793852>
- [14] Cohen W.M., Merrill S.A. (2003), Patents in the Knowledge-Based Economy. Washington, DC: The National Academies Press.
- [15] Hall B.H., Jaffe A., Trajtenberg M. (2001), Market Value and Patent Citations: A First Look. Economics Department Working Paper E00-277. University of California.
- [16] Hall B.H., Jaffe A., Trajtenberg M. (2005), Market Value and Patent Citations, "RAND Journal of Economics", 36 (1), pp. 16-38.
- [17] OECD. (2009), Patent Statistics Manual. Paris: OECD.

- [18] Smith A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*. <http://www.econlib.org/library/Smith/smWN1.html#nn1>
- [19] Say J.-B. (1803), *Traité d'économie politique ou simple exposition de la manière dont se forment, se distribuent et se composent les richesses*. Ire Editions. Crapelet.
- [20] de Sismondi S.J.C.L. (1819), *Nouveaux principes d'économie politique ou de la richesse dans ses rapports avec la population*. Paris: Delauney.
- [21] Lotz J.F.E. (1822), *Handbuch der Staatswirtschaftslehre*. Erlangen: Palm & Enke.
- [22] Jakob L.H. (1837), *Grundsätze der Polizeigesetzgebung und der Polizei—anstalten*. Halle: Grunert.
- [23] Mill J.S. (1848), *Principles of Political Economy with Some of their Applications to Social Philosophy*. <http://www.econlib.org/library/Mill/mlP.html>
- [24] Machlup F. (1958), *An Economic Review of the Patent System, Study of the Subcommittee on Patents, Trademarks, and Copyrights of the Committee on the Judiciary US Senate, 85th Congress, 2nd Session, Study No 15*. Washington, DC: US Government Printing Office. https://mises.org/sites/default/files/An%20Economic%20Review%20of%20the%20Patent%20System_Vol_3_3.pdf
- [25] Fisher I. (1912), *Elementary Principles of Economics*. New York: Macmillan.
- [26] Marshall A. (1919), *Industry and Trade: A Study of Industrial Technique and Business Organization*. London: Macmillan.
- [27] Vaughan F.L. (1925), *Economics of Our Patent System*. New York: Macmillan.
- [28] Clark J.B. (1927), *Essentials of Economic Theory*. New York: Macmillan.
- [29] Plant A. (1934), *The Economic Theory Concerning Patents for Inventions, "Economica"*, I, pp. 30-51.
- [30] Robbins L. (1939), *The Economic Basis of Class Conflict*. London: Macmillan.
- [31] Hayek F.A. (1944), *The Road to Serfdom*. London: Routledge.
- [32] Nordhaus W.D. (1972), *The Optimum Life of a Patent: Reply, "American Economic Review"*, 62, pp. 428-431.
- [33] Scherer F.M. (1972), *Nordhaus' Theory of Optimal Patent Life: A Geometric Reinterpretation, "American Economic Review"*, 62, pp. 422-427.
- [34] Horstmann I, Macdonald G.M., Slivinski A. (1985), *Patents as Information Transfer Mechanisms: To Patent or (Maybe) Not to Patent, "The Journal of Political Economy"*, 93, pp. 837-858.
- [35] Baumol W. (1990), *Entrepreneurship: Productive, Unproductive, and Destructive, "The Journal of Political Economy"*, 98, pp. 893-921.

- [36] Gilbert R., Shapiro C. (1990), Optimal Patent Length and Breadth, "The RAND Journal of Economics", 21 (1), pp. 106-112.
- [37] Klemperer P. (1990), How Broad Should the Scope of Patent Protection Be?, "The RAND Journal of Economics", 21 (1), pp. 113-130.
- [38] Cohen W.M., Nelson R.R., Walsh J.P. (2000), Protecting Their Intellectual Assets: Appropriability Conditions and Why US Manufacturing Firms Patent (or Not). NBER Working Paper. <http://www.nber.org/papers/w7552>
- [39] Stiglitz J. (2006), Give Prizes Not Patents, "New Scientist", Retrieved from <http://keionline.org/misc-docs/giveprizesnotpatents.pdf>, Retrieved date: 2017-01-15.
- [40] National Science Board (2010). Science and Engineering Indicators 2010. Arlington, VA: National Science Foundation (NSB 10-01). Retrieved from: <https://wayback.archive-it.org/5902/20150818153806/http://www.nsf.gov/statistics/seind10/pdf/seind10.pdf>, Retrieved date: 2017-01-15.
- [41] Scheu M., Veeffkind V., Verbandt Y., et al. (2006), Mapping Nanotechnology Patents: The EPO Approach, "World Patent Information", 28, pp. 204-211.
- [42] Magerman T., van Looy B., Song X. (2006), Data Production Methods for Harmonized Patent Statistics: Patentee Name Harmonization. KUL Working Paper <http://ec.europa.eu/eurostat/documents/3888793/5836029/KS-AV-06-002-EN.PDF>
- [43] Jaffe A.B., Fogarty M.S., Banks B.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, pp. 183-205.
- [44] Lanjouw J.O., Pakes A., Putnam J. (1998), How to Count Patents and Value Intellectual Property: Uses of Patent Renewal and Application Data, "The Journal of Industrial Economics", XLVI (4), pp. 405-433.
- [45] Johnson D. (2002), The OECD Technology Concordance (OTC): Patents by Industry of Manufacture and Sector of Use, OECD Science, Technology and Industry Working Paper 2002/5, Directorate for Science, Technology and Industry. Paris: OECD. Retrieved from [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/DOC\(2002\)5&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/DOC(2002)5&docLanguage=En), Retrieved date: 2017-01-15.
- [46] Popp D. (2005), Lessons from Patents: Using Patents to Measure Technological Change in Environmental Models, "Ecological Economics", 54, pp. 209-226.
- [47] Schmookler J. (1952), The Changing Efficiency of the American Economy: 1869-1938, "The Review of Economic Statistics", 34 (3), pp. 214-321.
- [48] Schmookler J. (1966), Invention and Economic Growth. (1st ed.). United States of America: Harvard University Press.
- [49] Basberg B. (1987), Patents and the Measurement of Technological Change: A Survey of the Literature, "Research Policy", 16, pp. 131-141.

- [50] Schankerman, M. (1998), "How Valuable Is Patent Protection? Estimates by Technology Field", *RAND Journal of Economics*, Vol. 29 (1), pp. 77-107, The RAND Corporation.
- [51] Pakes A., Griliches Z. (1980), Patents and R&D at the Firm Level: A First Look, "Economic Letters", 5, pp. 377-381.
- [52] Pakes A., Griliches Z. (1984), Patents and R&D at the Firm Level: A First Look. In Z. Griliches, editor. *R & D, Patents, and Productivity*. Chicago: University of Chicago Press, pp. 55-72.
- [53] Hausman J., Hall B.H., Griliches Z. (1984), Econometric Models for Count Data with an Application to the Patents-R&D Relationship, "Econometrica, Econometric Society", 52 (4), pp. 909-938. Retrieved from http://www.jstor.org/stable/1911191?seq=2#page_scan_tab_contents, Retrieved date: 2017-01-15.
- [54] Hall B.H., Griliches Z., Hausman, J.A. (1986), Patents and R and D: Is There a Lad?, "International Economic Review", 27 (2), pp. 265-286. Retrieved from http://www.jstor.org/stable/2526504?seq=17#page_scan_tab_contents, Retrieved date: 2017-01-15.
- [55] Pakes A. (1985), On patents, R&D, and the Stock Market Rate of Return, "Journal of Political Economy", 93, pp. 390-409. Retrieved from http://www.jstor.org/stable/1832183?seq=1#page_scan_tab_contents, Retrieved date: 2017-01-15.
- [56] Huang C. (2012), Estimates of the Value of Patent Rights in China. UNU-MERIT Working Papers. Maastricht: United Nations University-MERIT and Maastricht University.

INTECH