Patent data - The performance of firms, regions and countries

Patents can be used to monitor the technological performance of firms (or other organisations), regions or countries. Patents help track trends in technological leadership in a given technology field, e.g. by computing indices of revealed technological advantage. Compared to other output indicators, such as publications, patents are an indicator of activities closer to technology development. Matching patent information with additional information at the firm level allows the technology or patenting strategy of companies to be related to other characteristics. However, the attribution of a patent to a particular entity is not straightforward, and the quality of the attribution will have impacts on the quality of indicators generated.

Characteristics of patent data for analysing the technological performance of firms, regions and countries

The sectoral specialization of countries

The identification of technology domains and industries in patent data makes it possible to analyse the technological position of a country relative to others or to the world average. More specifically, the sectoral structure of countries' patenting activity can be investigated using patent indicators of specialisation (Soete and Wyatt, 1983). The most frequently used indicator is called the "specialisation index" or the "revealed technological advantage" (RTA) index, which is defined as the share of a country in patents in a particular field of technology divided by the country's share in all patents.

The index is equal to zero when the country holds no patents in a given sector. It is equal to 1 when the country's share in the sector is equal to its share in all fields (no specialisation), and grows (the upper limit will depend on the world distribution used) when a positive specialisation is found. The logarithm of the index can be used to obtain a new indicator with a distribution ranging from –1 to +1. Figures based on RTA indicators must be interpreted with caution, especially for international comparisons. A country with a very large total patent output will tend to have all its RTAs in the neighbourhood of 1, whereas a country with a low output of patents will have a very high value for the fields in which its output is slightly higher than the average for the country.
Specialisation indicators can be calculated for different periods, to show how countries' specialisation patterns have evolved over time. It should be remembered, however, that such indicators are relative to the world sectoral distribution of patents; if one country holds its distribution of patents steady while others increase their activity in an emerging field, its specialisation index in that field will decline.

Patents by companies

Information is provided in patent data for attributing a patent to a particular entity. The name and address of the patent holder are published in patent documents: however, the attribution of a patent to a particular entity is not so simple. Errors emerge because of spelling mistakes; companies have a number of different names (e.g. acronyms: IBM, International Business Machines); qualifications can be added to the name (e.g. Siemens, Siemens AG); patents can be taken by affiliates, some of which are easily identified (e.g. Sony US is an affiliate of Sony), whereas others are more difficult (Citroen is part of the PSA group). It is not unusual for a large group to have an affiliate in charge of managing its intellectual property, and the affiliate may file in its own name many of the group's patents (e.g. Philips).

Changes in the company's legal status, as well as changes in company names, affiliations, and mergers and acquisitions make the use of patentees' names in patent data an imperfect way to analyse company patenting and questions related to companies' patenting and innovation strategies. For instance, when aiming to harmonize a legal entity, all patents held by Hewlett Packard, Digital Equipment Corporation and Compaq might be considered as belonging to one and

the same legal entity; likewise, “Andersen Consulting” would become harmonised to “Accenture” (name change).

Methodological approaches can be commonly implemented for attributing a patent to a particular entity. Patent offices do some cleaning and harmonisation of names themselves. For instance, the USPTO deals with the name of the first applicant for any patent. The EPO attributes a standardised code to patent applicants, as does the JPO for applicants filing electronically. However, this is not sufficient to address the needs of statisticians. The cleaning and harmonisation of names may go through several steps (not all necessary or exclusive of each other) as follows:

- Basic cleaning (standardising abbreviations such as “Ltd”, “GmbH”, etc.) and standardisation of names.
- Matching the standardised name of applicants with a company database of reference (e.g. Amadeus for Europe, Compustat for the United States). Reconstructing the group structure by using information on the ownership structure (including affiliates) as reported in specialised databases (e.g. the “Who owns whom” database).

The first stage is identifying spelling variations to clean the names of applicants to obtain a standardised name for grouping companies. This is done with the aid of approximate matching techniques. Two approaches are used to group similar names and standardise. The rule-based approach involves the definition of rules to compare the similarity of names. The second approach relies on the use of dictionaries, large collections of names which serve as examples for a specific entity class (e.g. USPTO and EPO standard assignee names file; Derwent Patentee Codes). It is also possible for researchers to to build their own dictionaries with a harmonisation procedure (e.g. Magerman, Van Looy and Song 2006).

The second stage is to link the standardised names to the names contained in a company database (e.g. Amadeus, Compustat, etc.) directly or in combination with other methods to find as many potential matches as possible. For instance, other available information about the company (in addition to the name) can be used, e.g. addresses and searches based on related patentee names of priority patent filings or PCT (Patent Co-operation Treaty) applications. The matches obtained need to be validated, and doubtful matches can only be solved by hand. Lastly, the companies identified can be legally consolidated using information on the ownership structure. The two stages of matching and legal consolidation can also be carried out at the same time if the company data used already include information on the legal relationships between companies. However, data on ownership structure are rarely codified over time. As a result, most of the available information records only the most recent legal structure of companies. Consequently, further information is needed to track changes (e.g. mergers and acquisitions) over time and properly separate patenting activity by companies in different periods of time.

Major work in this field includes the NBER database of USPTO patents, harmonised with Compustat (www.nber.org/patents), the KUL algorithms for Eurostat (Magerman, Van Looy and Song, 2006), and the work done by Thoma and Torrisi (2007) and Thoma et al. (forthcoming).

Policy questions and patent indicators

Types of policy questions that can be addressed by patent data

- **The sectoral specialisation of countries.** As indicators of technological performance, the level of technological specialisation and/or strength of a geographical region or country (or company) can help policy makers identify weak and strong areas in national or regional innovation systems.

- **Patents by companies**. Matching patent information with other information at the firm level, such as R&D, innovation, stock market value, etc., makes it possible to relate the technology or patenting strategy of companies to other characteristics and respond to questions such as: What is the impact of patents on market value? What is the efficiency of R&D (in terms of patent numbers)? Moreover, attributing a patent to particular entities that own them is a key step in much statistical and analytical work based on patents. It allows the patent portfolio of companies to be reconstructed, which can be used to:

(1) Compile classifications of patents by industry, technical field, region, institutional sector, etc. and (2) Analyse the patenting strategy of firms (timing and orientation of their patent filings, in relation to competitors).

Technology specialization indicators

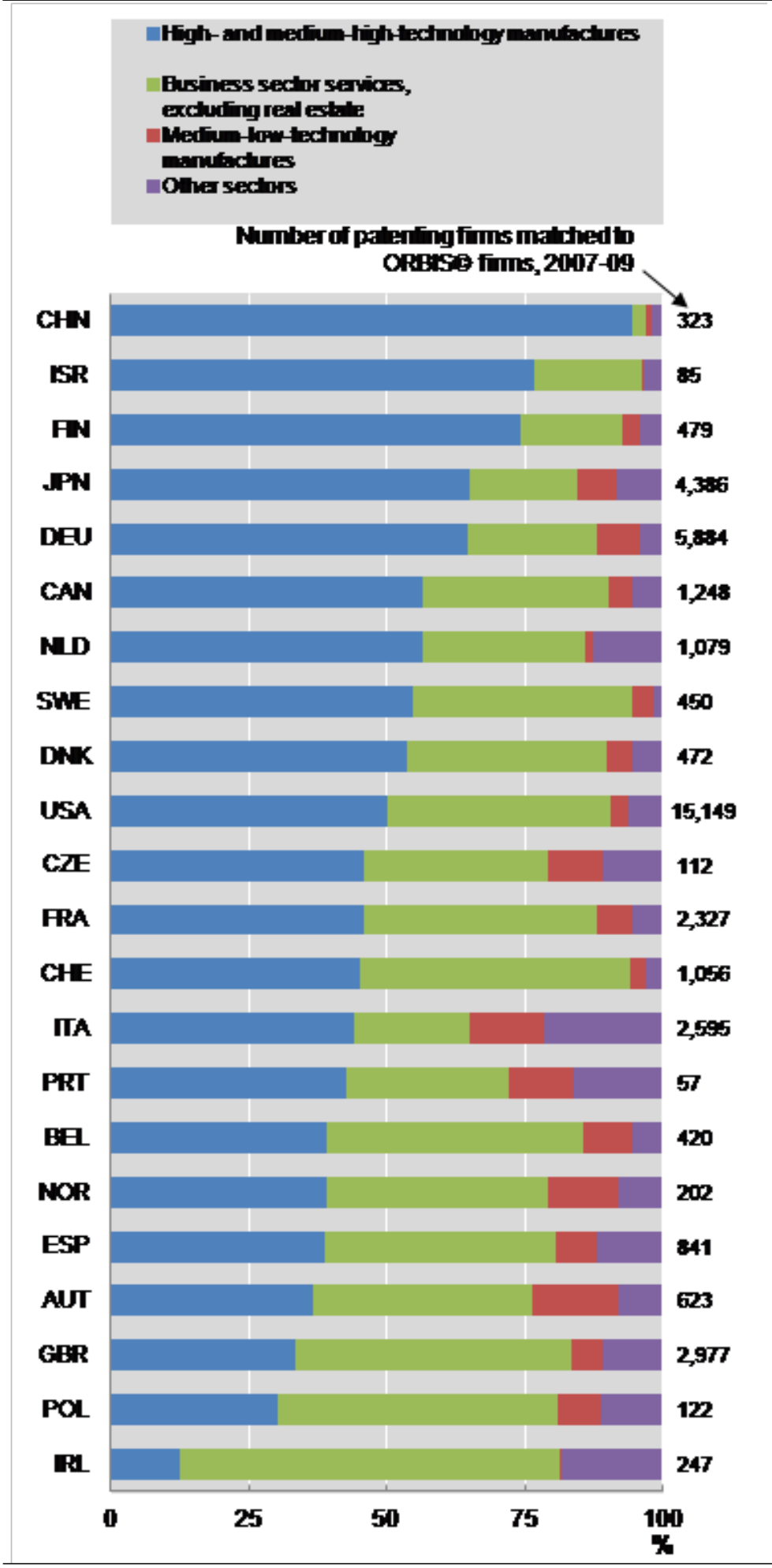
The revealed technological advantage index is based on patent counts and provides an indication of the relative specialization of a given country in selected technological domains.

The information provided by the International Patent Classification (IPC) constitutes a first reference for identifying patents in a specific domain. One or several IPC codes is attributed to the patent during the examination process. However, for emerging or enabling technologies, the patent classification system may not have a specific class. The OECD has designed definitions of ICT and biotechnology patents consisting of a list of IPC classes (www.oecd.org/sti/ipr-statistics [1]). The definitions, like the technologies, can evolve over time. This has been the case for nanotechnologies: in 2003 the EPO created a working group (NTWG) to develop a definition of the field that would identify nanotechnology patents through keyword searches and expert analysis. Patent applications from 15 countries or organisations were analysed and documents tagged as Y01N.

Patenting firms

Applicants in patent documents can be enterprises, organisations or persons. Business registers include information on enterprises and their main characteristics. By matching patent applicants' names to enterprise names in business registers, the patenting behaviour of firms can be linked to firm characteristics such as industrial sector, age and size. Matched patent and firm data can also review industries' contribution to the development of key technologies, such as biotechnology and information and communication technology (ICT).

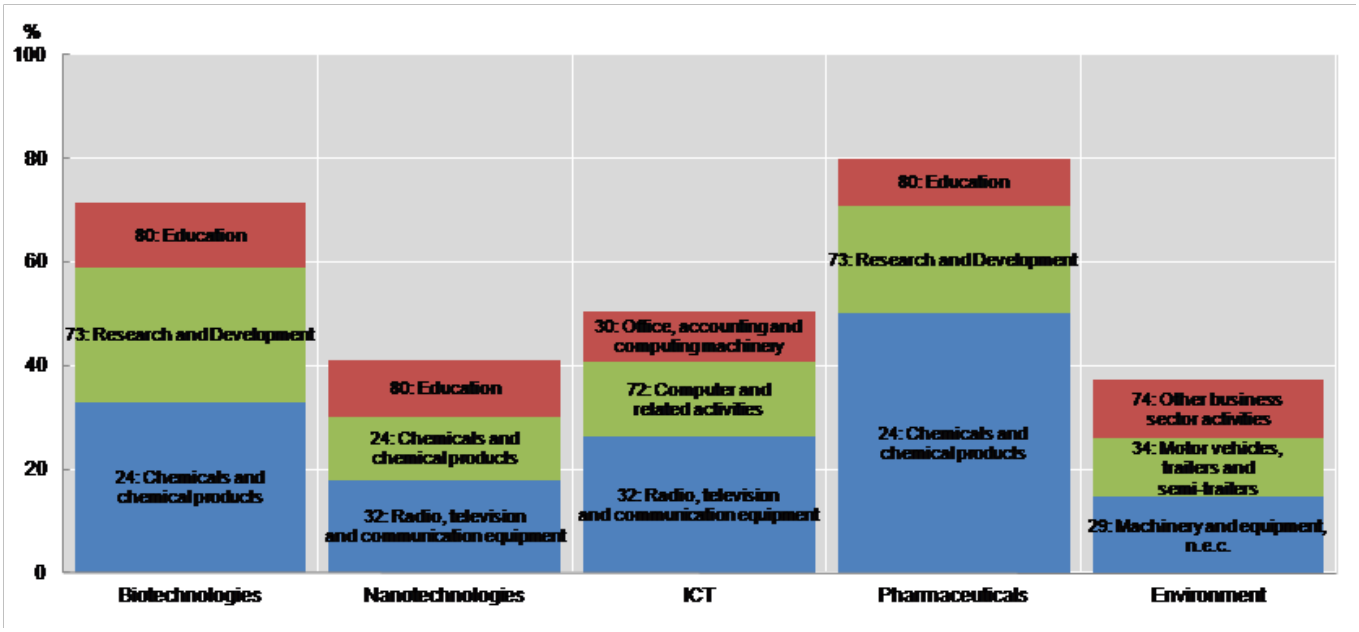
Figure 4. Patenting activity by sector, 2007–09, as % of patents filed at EPOa and USPTOb


High- and medium-high-technology manufactures
Business sector services, excluding real estate
Medium-low-technology manufactures
Other sectors
Number of patenting firms matched to ORBIS® firms, 2007-09
CHN 323
ISR 85
FIN 479
JPN 4,386
DEU 5,884
CAN 1,248
NLD 1,079
SWE 450
DNK 472
USA 15,149
CZE 112
FRA 2,327
CHE 1,056
ITA 2,595
PRT 57
BEL 420
NOR 202
ESP 841
AUT 623
GBR 2,977
POL 122
IRL 247
0
25
50
75
100
%

OECD Science, Technology and Industry Scoreboard 2011 — © OECD 2011
Source: OECD calculations based on the Worldwide Patent Statistical Database, EPO, April 2011; and ORBIS© Database, Bureau van Dijk Electronic Publishing, December 2010; matched using algorithms in the Imalinker system developed for the OECD by IDENER, Seville, 2011.
Notes: **a.** EPO= European Patent Office;
b. USPTO = U.S. Patent and Trademark Office.
c. Patenting firms were linked to the ORBIS© database using combinations of string matching algorithms that maximise the precision of the match. The patent portfolio of firms refers to patents applied at the European Patent Office (EPO) and at the U.S. Patent and Trademark Office (USPTO) between 2007 and 2009. Only countries with matching rates above 60% of patent filings over the period considered are shown.
The list of industries follows the ISIC (International Standard Industrial Classification), revision 3. High and medium-high-technology manufactures cover sectors 24, 29-35 less 351; medium-low technology manufactures include 23, 25-28, 351; and business sector services—excluding real estate—refer to 50-67, 71-74.
d. For Israel: "The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law. … It should be noted that statistical data on Israeli patents and trademarks are supplied by the patent and trademark offices of the relevant countries."

Matched enterprise and patent data also reveal the broad industrial basis of key enabling technologies.

Figure 5. Top three industries patenting in selected technology fields, 2007-09
Share of industries' contribution to patent applications in selected technology fields, EPO and USPTO patents



Source: OECD (2011), "OECD Science, Technology and Industry Scoreboard 2011", www.oecd-ilibrary.org/sites/sti_scoreboard-2011-en/06/11/index.html?cont... [2] OECD calculations based on the Worldwide Patent Statistical Database, EPO, April 2011; and ORBIS© Database, Bureau van Dijk Electronic Publishing, December 2010; matched using algorithms in the Imalinker system developed for the OECD by IDENER, Seville, 2011.
1. Patenting firms were linked to the ORBIS© database using combinations of string matching algorithms that maximise the precision of the match. The patent portfolio of firms refers to patents applied at the European Patent Office (EPO) and at the U.S. Patent and Trademark Office (USPTO) 2007–2009. Only countries with matching rates above 60% of patent filings over the period considered are included.
2. The list of industries follows the ISIC (International Standard Industrial Classification), revision 3.

3. Patents in biotechnologies and health- and ICT-related technologies are based on a selection of International Patent Classification (IPC) classes.
4. Patents in environment-related technologies are defined using combinations of IPC classes (for EPO and USPTO) and codes Y02 of the European Classification (ECLA) for EPO only.
5. Patents in nanotechnologies are identified by the ECLA code Y01.

The presence of young firms among patent applicants underlines the inventive dynamics of firms early in their development and their desire to develop new activities and products. This may affect their survival and growth.

Figure 1. Regional patents Hirshman-Herfindal index

References

- Magerman, T., B. Van Looy and X. Song (2006), “Data production methods for harmonized patent statistics: Patentee name harmonization”, KUL Working Paper No. MSI 0605.
- OECD (2013), “Innovation in science, technology and industry: OECD work on patent statistics”, OECD, Paris. http://www.oecd.org/innovation/inno/oecdworkonpatentstatistics.htm [3]
- OECD (2011), OECD Science, Technology and Industry Scoreboard 2011, OECD Publishing, Paris. doi: 10.1787/sti_scoreboard-2011-en
- OECD (2009), OECD Patent Statistics Manual, 5.2, 5.3; 5.6, 6.5.2, OECD Publishing, Paris. http://dx.doi.org/%2010.1787/9789264056442-en [4]
- Soete, L. and S. Wyatt (1983), “The use of foreign patenting as an internationally comparable science and technology output indicator”, Scientometrics No. 5/January, pp. 31–54.
- Thoma, G. L. D. and S. Torrisi (2007), “Creating powerful indicators for innovation studies with approximate matching algorithms: A test based on PATSTAT and Amadeus databases”, CESPRI Working Papers 211, CESPRI, Centre for Research on Innovation and Internationalisation, Universita’ Bocconi, Milan, Italy, revised December 2007.
- Thoma, G. L. D., S. Torrisi, A. Gambardella, D. Guellec, B. H. Hall and D. Harhoff (forthcoming), “Harmonisation of applicants’ names in patent data”, OECD Science, Technology and Industry Working Papers, Directorate for Science, Technology and Industry, OECD, Paris. www.oecd.org/sti/working-papers [5].

Source URL: https://www.innovationpolicyplatform.org/content/patent-data-performance-firms-regions-and-countries

Links
[1] http://www.oecd.org/sti/ipr-statistics
[2] http://www.oecd-ilibrary.org/sites/sti_scoreboard-2011-en/06/11/index.html?contentType=&i

temId=/content/chapter/sti_scoreboard-2011-65-en&containerItemId=/content/serial/20725345&accessItemIds=/content/book/sti_scoreboard-2011-en&mimeType=text/html;
[3] http://www.oecd.org/innovation/inno/oecdworkonpatentstatistics.htm
[4] http://dx.doi.org/%2010.1787/9789264056442-en
[5] http://www.oecd.org/sti/working-papers