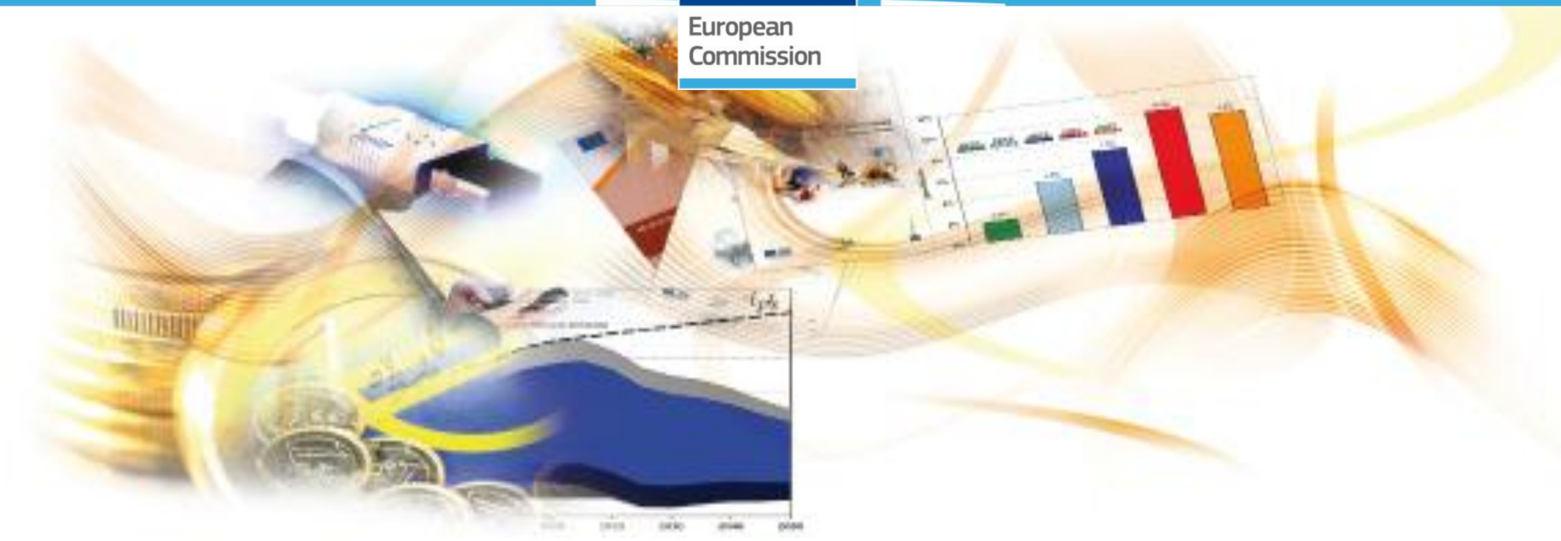


JRC TECHNICAL REPORTS

The patenting activity of the top IRI Scoreboard Companies: an introductory note

Antonio Vezzani, Fabio Montobbio, Sandro Montresor and Gianluca Tarasconi

2014

Report EUR 26350 EN

Joint Research Centre

European Commission
Joint Research Centre
Institute for Prospective Technological Studies

The main authors of this paper are: Antonio Vezzani (European Commission, JRC-IPTS, Knowledge for Growth Unit, Economics of Industrial Research and Innovation Action), **Fabio Montobbio** (Dep. of Economic and Statistics 'Cognetti de Martiis', University of Turin and CRIOS - Center for Research in Innovation, Organization and Strategy, Bocconi University), **Sandro Montresor** (European Commission, JRC-IPTS, Knowledge for Growth Unit, Economics of Industrial Research and Innovation Action and University of Bologna), and **Gianluca Tarasconi** (CRIOS - Center for Research in Innovation, Organization and Strategy, Bocconi University and OST - Observatoire des sciences et des techniques, Paris).

Acknowledgement: The authors are grateful to Fernando Hervás for his comments on a previous version of this work, usual caveats apply.

Contact information
Fernando Hervás Soriano
Address: European Commission, Joint Research Centre - Institute for Prospective Technological Studies
Edificio Expo. C/ Inca Garcilaso, 3 E-41092 Seville (Spain)
E-mail: jrc-ipts-kfg-secretariat@ec.europa.eu
Tel.: +34 95 448 84 63
Fax: +34 95 448 83 26
IPTS website: http://ipts.jrc.ec.europa.eu/; JRC website: http://www.jrc.ec.europa.eu

JRC 86166

EUR 26350 EN

ISBN 978-92-79-34744-3 (pdf)

ISSN 1831-9424 (online)

doi:10.2791/53758

Luxembourg: Publications Office of the European Union, 2014

Printed in Spain

Abstract

This note presents the main results of an explorative and introductory analysis of the patenting activity exhibited by the top 100 companies of the IRI Scoreboard. The main purpose of this pilot project has been to identify strengths and weaknesses of the exercise, in order to consider its possible future extension to the whole Scoreboard sample. With respect to these top 100 companies, patent families are built up on the basis of patent data drawn from Patstat and results are analysed by comparing them with other data related to their R&D investments. We observe that both the R&D and the patent applications of the investigated sample of companies increase over time. At the same time, important sector specificities in the R&D-patent relationship have been found. The analysis of the technological competences of the overall sample yields promising results. A first examination of the IPC classes of the patent applications suggests a certain concentration in the kind of technological knowledge that companies master. The analysis of the knowledge base and, more specifically, the companies' involvement in the creation of key enabling technologies (KETs) also highlights that important sector specificities go along with firm-specific factors. All-in-all "augmenting" the Scoreboard data with company level patent information appears to be an interesting extension to be pursued.

Keywords: patents, technological profile, KETS, R&D

JEL codes: O30, O31, O32

1 Introduction

This note contains an explorative, introductory analysis of the patent data provided to the IRI team by the CRIOS (Center for Research in Innovation, Organization and Strategy) research centre.[1] This has been undertaken on the basis of a scientific collaboration between individual researchers of the JRC-IPTS (Antonio Vezzani and Sandro Montresor) and CRIOS (Fabio Montobbio and Gianluca Tarasconi). The objective has been to approach the analysis of the patenting activity of the EU Industrial R&D Investment Scoreboard companies,[2] as part of the research

[1] CRIOS is an inter-departmental research centre, established at the premises of the Bocconi University of Milan (see http://www.crios.unibocconi.eu for further details).
[2] See http://iri.jrc.ec.europa.eu/scoreboard.html

activities included in the IRIMA[3] project (Research Topic: "Technological profile and innovation patterns of European top R&D investors").

One of the objectives of the IRI action is to define and analyse the technological profiles of the top world R&D investors, to better understand the current and prospective capacity of European companies to compete globally in a series of key strategic high R&D sectors. In fact, side by side with a large population of SMEs, Europe can rely on a relatively large number of multinational companies with very high R&D expenditures (ranked in the IRI Scoreboard). Their patenting activities require careful attention, as this enables us to analyse important opportunities for technological diversification and competition.

This note presents the first results of a pilot study of the patent portfolios of the top 100 Scoreboard companies. With respect to these 100 companies, patent data have been built up by CRIOS using the information on their ownership structure and the set of “rules” illustrated in Technical Annex I. Following current practice in the field, the retained patents are only those filed by the inventors at one or more of the main “international” patent offices (EPO – USPTO – WO) during the period 2000-2010. Moreover, in order to control for the multiple patenting of the same invention in different countries, we have matched the patent applications with the INPADOC families' data (for a definition of the INPADOC family, see Annex III). Accordingly, hereafter in the present document one patent is meant as one patent family, that is: a set of patents that refer to the same priority date (first application in a given office). The note reports and comments on a first set of descriptive statistics that mainly refer to the 100 companies as a whole. In so doing, a number of aspects are identified, which could/should represent the subject of further investigation of data at the individual company level.

The note is structured in the following way. Section 2 gives a first glance at the patent data and an example of aggregated combined analysis with R&D investment data. Section 3 presents an analysis of the companies’ technological competences. Section 4 presents the top 100 IRI Scoreboard companies' patent portfolios and sketches their technological profiles and engagement in the so-called “Key Enabling Technologies” (KETs). Section 5 concludes.

[3] The Industrial Research and Innovation Monitoring and Analysis is a project implemented under a collaborative agreement between the JRC and DG RTD.

2 A first glance

Over the retained period, the top 100 R&D investors of the IRI Scoreboard filed about 25% of the total patents recorded in Patstat[4] (world patent applications). Bearing in mind that the same companies account for nearly 54% of the total Business R&D of OECD countries,[5] it emerges that the extent to which their prominent innovative efforts translate into a prominent innovative outcome is limited. The linear model of innovation ("more input, more output") does not seem to apply suggesting us to consider, also with respect to these top innovative investors, the uncertain nature of their R&D projects, the innovative role of other intangibles in addition to R&D (e.g. training, design, and the like), and alternative instruments to appropriate innovation outcomes than those represented by patents.

Figure 1 shows the trend in the number of patent applications over the period 2000-2010, along with that of (the logarithm of) R&D from the earliest available year of the IRI Scoreboard (2004).

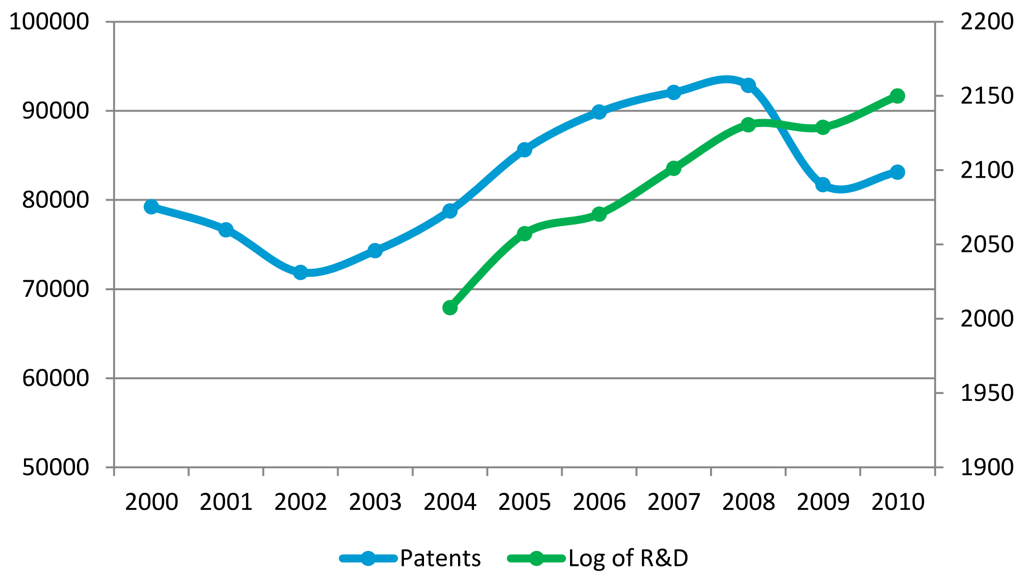


Figure 1: Patent applications and R&D expenditure over time

As usually occurs in this kind of analysis, the number of patent applications in the most recent years could suffer from a downward bias, due to the technical delay in their formal registration. However, a clear pattern emerges. Both patents and R&D of the 100 companies at stake have increased substantially in the last decade (the average rate of change is about 2.1% and 5.7%,

[4] The total number of patents has been obtained by selecting all the patents reported in the Patstat dataset with the same publication authority (EPO – USPTO – WO) of our sample.

[5] With a certain benefit of hindsight, this share has been obtained by relating the R&D figures of our sample for 2011 with BERD data from OECD in the same year.

respectively[6]), pointing to their successful commitment to stay at, and possibly push past, the innovation frontier. On the other hand, patents have increased in this period less rapidly than R&D, suggesting an apparent declining ratio of patent applications to R&D spending.[7] For these 100 companies, innovative efforts might have entered diminishing innovative returns: a hypothesis that deserves and requires closer scrutiny in future research.

A further step in the analysis is to look at what we could call the "R&D innovative efficiency" of these 100 companies, calculated as the ratio between their total number of patents filed and their total R&D expenditure (expressed in millions of euros) over the period 2004-2010. Given the specificities that the relationship between R&D and patents reveals (among others) at the industry level - one just needs to think about the famous distinction between Schumpeter Mark I and Mark II industries, for which see, for example, Breschi et al (2003) - a sector level analysis is relevant. The restricted size of the sample (100 companies) and the problems of representativeness that it could create in some sectors, do not allow us to retain the ICB 3-digit level at which Scoreboard companies are classified. Their aggregation into 8 "macro-sectors" appears to attenuate this problem and is used hereinafter.

Table 1 reports the distribution of the companies across ICB 3-digit Sectors; 57 out of the 100 companies operate in the "Pharmaceutical and biotechnology", "Technology hardware & equipment", or "Automobiles & parts" sectors.

Table 1: Distribution of companies across ICB-3digit sectors

ICB-3digit sector	Number of firms	Sector Group
Pharmaceuticals & biotechnology	21	Pharma & Health
Technology hardware & equipment	19	Tech. hardware
Automobiles & parts	17	Auto
Electronic & electrical equipment	7	Electronic
Software & computer services	7	Software
Chemicals	6	Medium
Aerospace & defence	4	Medium
General industrials	3	Medium
Leisure goods	3	Leisure
Banks	2	Low
Food producers	2	Low
Industrial engineering	2	Medium
Oil & gas producers	2	Low

[6] The average growth rate of patents has been calculated over the period 2000-2008, whereas that of R&D considers the period 2004-2010.
[7] Similar evidence is found, on US data, by Daniel Wilson (2003).

Fixed line telecommunications	1	Low
General retailers	1	Low
Health care equipment & services	1	Pharma & Health
Household goods & home construction	1	Medium
Mining	1	Low

In addition, the three firms operating in the "Leisure" sector are among the top patentees, with a number of applications well above the sample average. For these reasons in presenting the sectorial statistics, companies will be grouped according to 8 "macro-sectors": 1) Leisure; 2) Electronic; 3) Software; 4) Technology hardware; 5) Automobiles and parts; 6) Other companies operating in sectors with medium R&D intensity (Medium); 7) Pharmaceuticals & biotechnology plus Health care equipment & services (Pharma & Health); 8) Companies operating in low R&D intensity sectors (Low).

As Figure 2 shows, the 7 companies operating in the Electronic & electrical equipment industry (that is, FUJIFILM, LG, Mitsubishi Electric, Renesas, Samsung Electronics, Sharp and Siemens) have the highest R&D efficiency in terms of patents, closely followed by those operating in the Leisure and Software industries. Conversely, companies operating in the "Pharma & Health" sector are among those with the lowest ratio. Although further analysis is required before drawing any conclusion, the R&D efficiency of the scoreboard companies appears to have a big component of sector specificity, which will have to be considered along with firm-specific factors.

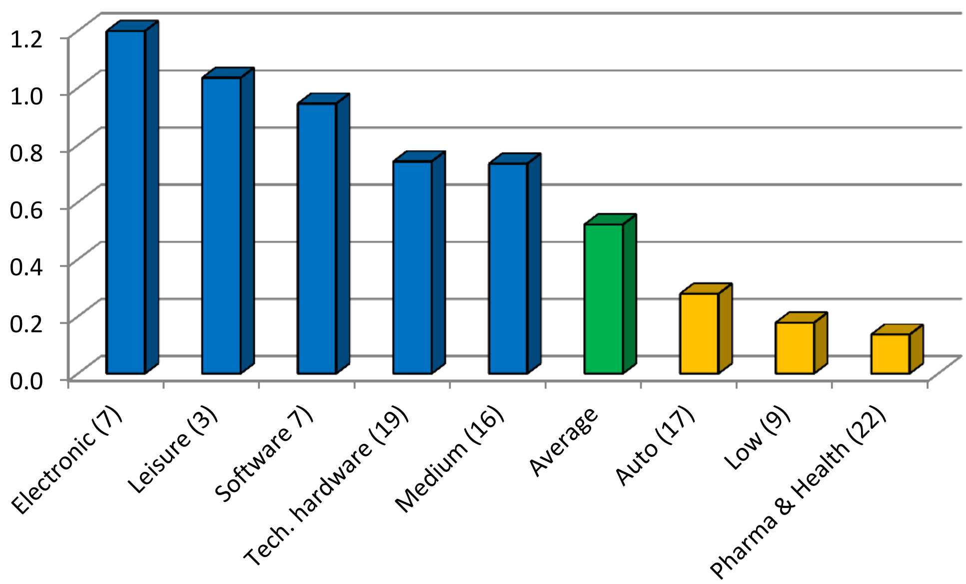


Figure 2: R&D innovative "efficiency" by sector group

3 The technological competencies of the (top 100) Scoreboard Companies

The matching between company and patent data allows us to disentangle the technological fields in which the investigated Scoreboard companies have obtained their inventions. Following the extant literature, we can take their portfolio of patent classes as a proxy of the array of technological competencies that they master.[8]

As usual, the reference to the IPC classification of the retained patents (for its hierarchical structure, see Technical Annex II) can be helpful in this last respect. A first exploration is reported in Figure 3, which shows the shares of IPC "sections" (the most aggregated level of the IPC classification) "activated" by the patents filed by our sample of 100 companies. Interesting insights emerge from this preliminary snapshot.

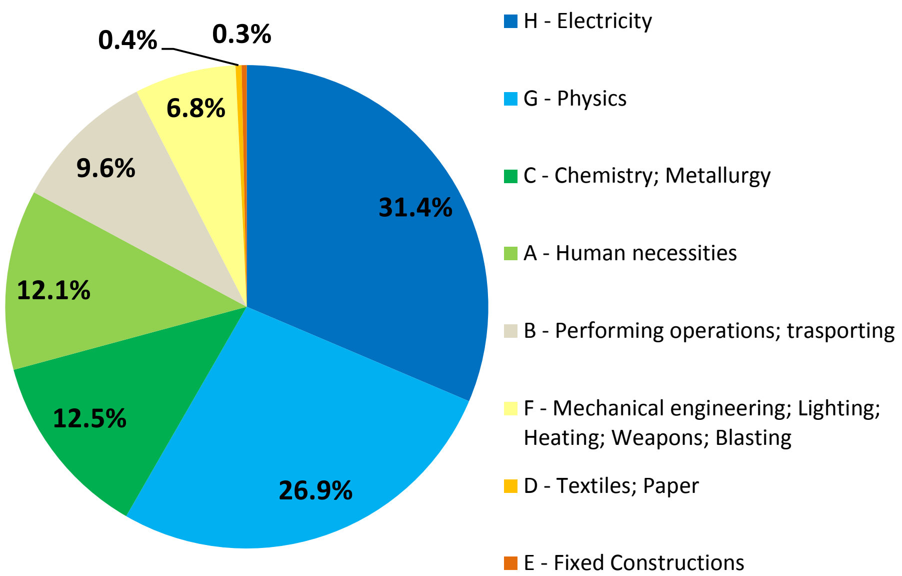


Figure 3: Patents by IPC section

From this aggregated picture, a significant degree of technological concentration appears. The subclasses[9] contained in the Electricity section are indicated by 31.4% of the patents of the sample, followed by Physics (26.9%). Chemistry and Metallurgy, and Human necessities show similarly remarkable shares (12.5% and 12.1% respectively), while the other sections stay at major distance. The top 100 Scoreboard R&D investors appear to share a quite focused technological base, at least in terms of the principles and knowledge elements of the four

[8] This is an idea that dates back at least to the seminal work by Pari Patel and Keith Pavitt (1997).

[9] In order to calculate the shares of IPC sections and subclasses, single entries of IPC 8-digit codes have been considered for each patent family.

sections above. This is an interesting result, whose validity is worthy of investigation with respect to the larger sample of Scoreboard R&D investors.

In order to have a more fine-grained picture, a similar analysis can be carried out by looking at the shares of IPC "subclasses" of the overall sample, a more disaggregated level of the IPC system (just to give an idea, the total number of technological subclasses attached to the patents analysed is more than 3,500,000). Figure 4 reports the 10 most reported subclasses, which are referred by about 43% of the patents of the 100 sampled companies.

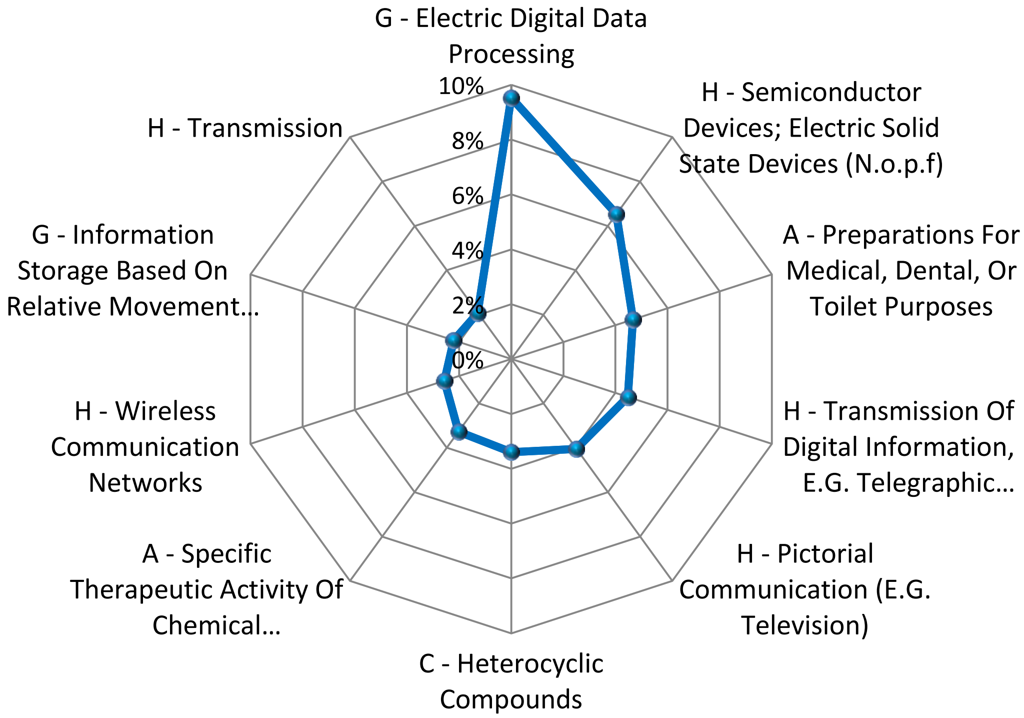


Figure 4: Patents by IPC subclass

The distribution across these subclasses is coherent with that observed across the corresponding sections, but appears *a fortiori* less concentrated. Still, in general, few subclasses represent the core of the corresponding sections, pointing again to a certain homogeneity of technological competencies across the 100 companies at stake. For example, "Human necessities" (section A) is mainly represented by two subclasses: "Preparations for medical, dental, or toilet purposes" and "Specific therapeutic activity of chemical compounds or medicinal preparations". These two subclasses alone account for about 2/3 of the whole corresponding section. Overall, traces of "not much variety"[10] seem to emerge also with respect to the Scoreboard companies: an issue that of course deserves and requires further investigation.

[10] A well-known article of Patel and Pavitt (1997) pointed out that patent portfolios are complex and path-dependent, but the technological competences do not vary drastically between firms of a given industry. A range of

4 Technological profiles and Key Enabling Technologies (KETS)

In this section we present a preliminary analysis of the patent portfolios for our sample of Scoreboard companies. Research questions to address would include: How do the top R&D investors rank in terms of patents (in our case, patent applications)? How do they build up their patent portfolios? What are their technological competences? Do R&D leaders (followers) stay ahead (lag behind) also in terms of patent applications? Does their R&D and patent ranking correlate with their relative position in the development of key technologies, like for example KETs? More generally, do country and sector specificities also emerge in a patent-augmented version of the IRI Scoreboard?

In order to respond to these questions, the patent portfolios of the top 100 were first classified following the Technology classification proposed by Schmoch (2008) and adopted by the WIPO. In a second step, in order to test a more focused analysis on a particular group of technologies, we classified patent portfolios following the classification contained in "Feasibility study for an EU Monitoring Mechanism on Key Enabling Technologies" (2012).

It should be noted that in completing the matching some practical problems arise. First of all, these classifications are not traceable in terms of patent codes to the same extent, with the risk of introducing some bias in their analysis. For example, with respect to the classification of KETs we need to compare Micro- and nano-electronics – which covers new technologies related to semiconductors – piezo- or nano-electronics, all easily identifiable through IPC classes, with Industrial biotechnology, which is more difficult to define since many classes covering inventions related to it may also cover new technologies linked to red or green biotechnology. Second, the codes identified for the two classifications refer to different levels of the IPC one (of those described in Technical Annex II), with another possible source of bias.[11] Last, but not least, the mapping is complicated by the fact that each and every patent normally refers to a number of different IPC classes. With this last respect in the present note we adopted a different approach when dealing with the two classifications adopted, as will be discussed in the following section.

competing and differentiated products may be based on the same technologies, thus, *product variety in an industry is compatible with technological homogeneity.*

[11] When reported to the Subgroup level, which is the lowest one in the hierarchy of the IPC classification, the KETs-related codes correspond to 4,361 different entries.

4.1 A first glance at the companies' technological profiles...

The technological classification proposed by Schmoch (2008) has been conceived in order to allow country comparisons and stresses the importance of technological competences as basis for engagement in specific product areas and sectors.

The classification is based on IPC codes and has a hierarchical structure. The first level, used in this note, is made of 5 broad categories (areas): 1) Electrical engineering; 2) Instruments; 3) Chemistry; 4) Mechanical engineering; 5) Other fields. In addition, a more detailed level of the classification that defines 35 subgroups (fields) allowing for more in-depth analyses could be used.

In order to outline the technological portfolio of the companies under study we used a fractional counting approach. In brief, each patent has been weighted by the ratio of the number the IPC codes belonging to a specific technological field x over the total number of IPC codes reported:

$$pw = \sum ipc\_in\_field\_x / \sum ipc\_classes \leq 1$$

When calculating the total number of patents in each technological area (and field) in order to define the companies' portfolio, each patent has been thus counted as *pw*.

Overall, the top 100 Scoreboard exhibit a quite specialised set of technological competences, mainly based on electrical engineering (see fig. 5). However, strong sector specificities also arise with this last respect. Companies operating in IT-related sectors show, on average, a patent portfolio very focused on electrical engineering competences. These are even more fundamental in the Software sector where they represent about 90% of the total patents. On the other hand, companies operating in Pharma & Health sectors are those with a set of competences more based on Chemistry knowledge (followed by those operating in the Low and Medium sector groups). As expected, companies operating in the Automobile sectors show strong competences related to mechanical engineering, whereas instruments competences seems to play a non-marginal role for almost all the companies analysed.

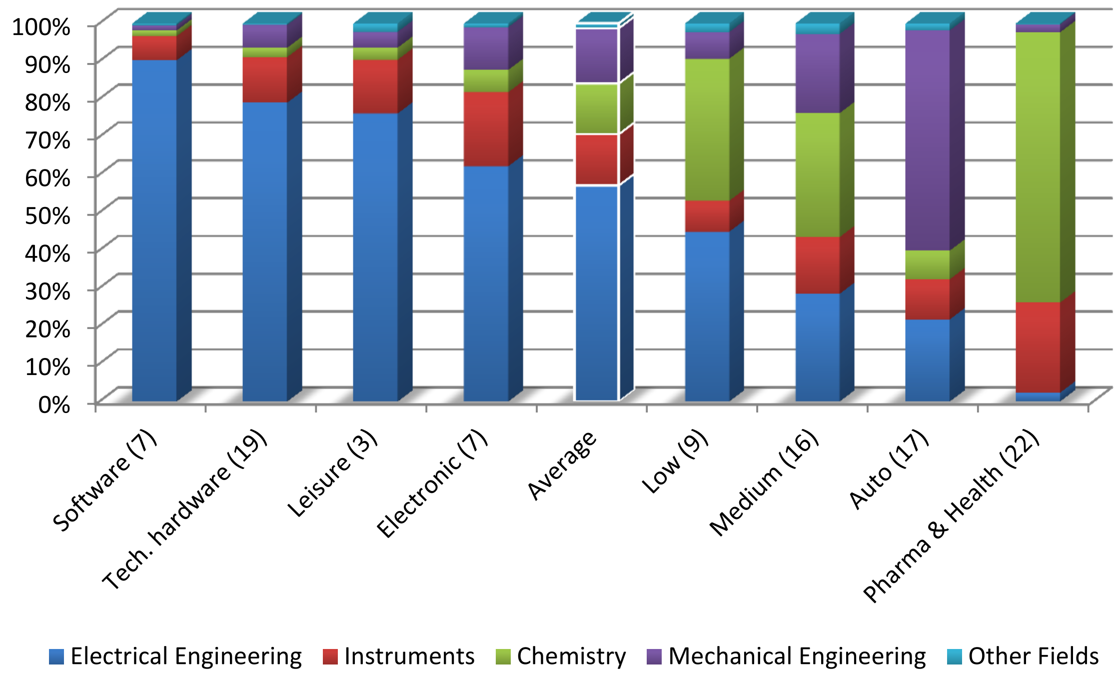


Figure 5: Patent distribution across technological areas by sector group

Table 2 reports the top 100 R&D investors ranked on the basis of the total number of patent applications during the period 2000-2010. Their patent distribution across the 5 broad technological areas presented above already gives some interesting insights (for each area the 10 highest shares are highlighted in green). With 91% of patents in the electrical engineering area, IBM, the top company in terms of patent applications, does not figure as one of the most specialised in this area (Huawei and ZTE have a share of patents above 98%). Overall, the companies' specialisation in the instrument area appears to be less strong, but 77% of firms have at least 5% of patents; Medtronic is the most specialised company with about 86% of the total patent portfolio. 9 out of 10 companies more specialised in chemistry operate in the Pharma & Health sector, being Monsanto, classified as Food Producer, the tenth company. Of course, chemical companies are highly specialised in this area, but with a non-marginal share of patents distributed among the others. The Automobile sector’s specialisation in mechanical engineering is confirmed at company level; all the companies with a high share of patents in this area operate in this sector, United Technologies (Aerospace & defence) being the only non-automobile company with a share above 50% (60%).

Table 2: Total patents and their distribution across technological areas of the top 100 Scoreboard companies

Company	ICB classification (3-digit)	Total Patents	Electrical Engineering (%)	Instruments (%)	Chemistry (%)	Mechanical Engineering (%)	Other Fields (%)
IBM	Software & computer services	50948	91.0%	5.9%	1.3%	1.4%	0.4%
Samsung Electronics	Electronic & electrical equipment	46792	76.9%	13.3%	3.3%	4.7%	1.8%
Panasonic	Leisure goods	40282	76.4%	11.7%	3.6%	6.1%	2.1%
Siemens	Electronic & electrical equipment	32713	52.9%	19.6%	7.1%	19.7%	0.7%
General Electric	General industrials	32450	31.1%	23.6%	16.1%	25.5%	3.7%
Canon	Technology hardware & equipment	28068	51.6%	27.2%	5.3%	15.5%	0.3%
Fujitsu	Software & computer services	26013	83.5%	10.5%	2.7%	2.9%	0.4%
Sony	Leisure goods	25960	83.3%	10.0%	2.2%	2.6%	1.9%
Toshiba	General industrials	25460	74.2%	14.3%	3.2%	7.4%	0.8%
Philips	Leisure goods	25008	68.4%	22.6%	3.6%	2.6%	2.8%
Hitachi	Technology hardware & equipment	24584	61.6%	16.4%	8.0%	13.2%	0.9%
Microsoft	Software & computer services	22682	96.4%	2.2%	0.1%	0.2%	1.1%
FUJIFILM	Electronic & electrical equipment	20582	36.4%	32.0%	15.6%	15.8%	0.3%
Hewlett-Packard	Technology hardware & equipment	19450	73.1%	9.6%	2.9%	13.4%	0.9%
Robert Bosch	Automobiles & parts	19080	24.1%	16.8%	4.9%	53.4%	0.8%
Toyota Motor	Automobiles & parts	16261	23.3%	7.2%	11.9%	56.6%	1.0%
Nokia	Technology hardware & equipment	15987	93.6%	5.0%	0.2%	0.5%	0.7%
Intel	Technology hardware & equipment	15392	88.7%	8.4%	1.4%	1.2%	0.2%
NEC	Software & computer services	15224	84.1%	10.3%	3.8%	1.5%	0.3%
Mitsubishi Electric	Electronic & electrical equipment	13631	61.6%	14.4%	2.3%	21.0%	0.8%
BASF	Chemicals	13517	2.8%	2.3%	85.7%	8.5%	0.7%
Sharp	Electronic & electrical equipment	13402	65.3%	25.2%	3.3%	5.1%	1.0%
Google	Software & computer services	12696	90.9%	6.0%	1.1%	1.4%	0.6%
Alcatel-Lucent	Technology hardware & equipment	12619	88.3%	8.9%	1.0%	1.5%	0.2%
Ericsson	Technology hardware & equipment	12261	95.9%	3.4%	0.1%	0.4%	0.2%
Honda Motor	Automobiles & parts	11784	19.2%	7.5%	8.4%	62.4%	2.6%
Honeywell	General industrials	11121	31.6%	26.4%	19.9%	20.1%	2.1%
Denso	Automobiles & parts	10387	35.3%	18.5%	5.1%	40.1%	1.0%
Texas Instruments	Technology hardware & equipment	10252	83.6%	13.8%	1.1%	1.3%	0.2%
General Motors	Automobiles & parts	10146	25.4%	8.8%	8.8%	55.2%	1.8%
Oracle	Software & computer services	9294	95.5%	3.5%	0.1%	0.8%	0.2%
Qualcomm	Technology hardware & equipment	9017	92.2%	6.7%	0.6%	0.3%	0.3%
DuPont	Chemicals	9012	11.8%	4.7%	70.0%	11.7%	1.8%
Ricoh	Technology hardware & equipment	8975	43.3%	38.1%	2.9%	15.3%	0.4%
Bayer	Chemicals	8656	2.5%	12.1%	80.0%	5.0%	0.4%
Huawei	Technology hardware & equipment	8620	98.8%	1.0%	0.0%	0.2%	0.1%
United Technologies	Aerospace & defence	8250	15.7%	8.4%	11.3%	60.1%	4.6%
Cisco Systems	Technology hardware & equipment	7120	95.6%	4.0%	0.0%	0.2%	0.1%
Procter & Gamble	Household goods & home construction	7062	7.4%	19.5%	47.0%	18.5%	7.6%
STMicroelectronics	Technology hardware & equipment	6467	87.5%	9.7%	1.6%	1.0%	0.2%
Pfizer	Pharmaceuticals & biotechnology	6419	0.8%	5.3%	91.7%	2.0%	0.2%
Research In Motion	Technology hardware & equipment	6029	95.5%	3.7%	0.1%	0.4%	0.3%
ZTE	Technology hardware & equipment	5984	98.9%	1.0%	0.0%	0.1%	0.0%
Boeing	Aerospace & defence	5892	34.5%	23.0%	6.9%	33.7%	1.8%
Delphi	Automobiles & parts	5878	28.5%	12.7%	6.0%	50.7%	2.1%
Ford Motor	Automobiles & parts	5819	15.5%	6.8%	8.9%	67.0%	1.8%
Medtronic	Health care equipment & services	5716	6.4%	85.6%	5.6%	2.3%	0.1%
Advanced Micro Devices	Technology hardware & equipment	5653	85.7%	11.3%	1.5%	1.4%	0.1%
Volkswagen	Automobiles & parts	5512	8.1%	6.7%	7.6%	74.9%	2.6%
Roche	Pharmaceuticals & biotechnology	5433	3.9%	26.3%	68.1%	1.7%	0.1%
LG	Electronic & electrical equipment	5415	54.5%	37.5%	4.2%	2.9%	0.9%
Renesas	Electronic & electrical equipment	5381	90.9%	7.4%	1.0%	0.7%	0.1%
Continental	Automobiles & parts	5127	22.3%	15.8%	4.2%	56.3%	1.4%

Dow Chemical	Chemicals	5109	6.8%	6.7%	73.2%	11.8%	1.5%
GlaxoSmithKline	Pharmaceuticals & biotechnology	4633	1.0%	7.0%	89.5%	1.9%	0.7%
Broadcom	Technology hardware & equipment	4552	95.4%	4.3%	0.0%	0.0%	0.1%
EADS	Aerospace & defence	4446	18.8%	18.1%	5.7%	53.9%	3.4%
Nissan Motor	Automobiles & parts	4442	23.6%	4.9%	9.7%	60.8%	1.0%
Novartis	Pharmaceuticals & biotechnology	4352	0.8%	17.3%	77.7%	4.0%	0.2%
Merck US	Pharmaceuticals & biotechnology	4329	1.2%	4.1%	93.5%	0.7%	0.5%
NTT	Fixed line telecommunications	4260	86.8%	10.7%	1.5%	0.7%	0.3%
Sumitomo Chemical	Chemicals	4173	18.8%	7.9%	67.3%	5.7%	0.3%
Johnson & Johnson	Pharmaceuticals & biotechnology	4161	2.4%	63.8%	29.7%	3.6%	0.5%
Abbott Laboratories	Pharmaceuticals & biotechnology	3930	3.5%	45.2%	47.1%	3.7%	0.6%
Aisin Seiki	Automobiles & parts	3857	15.9%	13.7%	4.1%	59.0%	7.3%
SAP	Software & computer services	3798	97.2%	2.5%	0.0%	0.2%	0.0%
Apple	Technology hardware & equipment	3553	91.5%	4.8%	0.7%	1.3%	1.8%
Mitsubishi Chemical	Chemicals	3540	13.3%	9.2%	66.9%	10.1%	0.5%
Caterpillar	Industrial engineering	3507	15.7%	7.6%	10.0%	56.3%	10.3%
Merck DE	Pharmaceuticals & biotechnology	3290	7.4%	9.7%	79.7%	2.8%	0.4%
Peugeot (PSA)	Automobiles & parts	3267	6.4%	3.7%	9.4%	75.9%	4.6%
Daimler	Automobiles & parts	2958	19.5%	7.8%	8.4%	62.1%	2.1%
Sanofi-Aventis	Pharmaceuticals & biotechnology	2896	0.7%	14.8%	83.0%	1.3%	0.1%
AstraZeneca	Pharmaceuticals & biotechnology	2795	0.4%	5.6%	92.7%	1.3%	0.0%
Renault	Automobiles & parts	2725	7.6%	4.8%	8.5%	77.3%	1.7%
Nestle	Food producers	2572	2.1%	4.1%	65.9%	20.5%	7.5%
BMW	Automobiles & parts	2511	13.8%	7.1%	5.5%	70.9%	2.7%
Volvo	Industrial engineering	2451	6.1%	5.9%	10.1%	56.8%	21.1%
Bristol-Myers Squibb	Pharmaceuticals & biotechnology	2445	0.6%	6.8%	91.3%	1.3%	0.1%
Hyundai Motor	Automobiles & parts	2435	15.1%	5.5%	9.2%	66.4%	3.7%
Monsanto	Food producers	2347	0.6%	1.3%	94.4%	3.7%	0.0%
EMC	Technology hardware & equipment	2138	97.1%	1.4%	0.3%	0.7%	0.5%
Boehringer Ingelheim	Pharmaceuticals & biotechnology	2006	0.5%	9.9%	86.4%	3.1%	0.1%
Eli Lilly	Pharmaceuticals & biotechnology	1259	0.6%	5.0%	93.7%	0.7%	0.1%
Royal Bank of Scotland	Banks	1148	61.7%	25.1%	2.8%	8.8%	1.6%
Fiat	Automobiles & parts	1098	15.7%	7.0%	3.9%	71.4%	2.0%
Daiichi Sankyo	Pharmaceuticals & biotechnology	1086	0.1%	1.5%	98.0%	0.5%	0.0%
Takeda Pharmaceutical	Pharmaceuticals & biotechnology	1072	1.0%	8.5%	89.9%	0.6%	0.0%
Novo Nordisk	Pharmaceuticals & biotechnology	1014	3.6%	26.3%	67.5%	2.5%	0.1%
Amgen	Pharmaceuticals & biotechnology	961	0.3%	5.0%	94.1%	0.6%	0.0%
Astellas Pharma	Pharmaceuticals & biotechnology	747	0.8%	5.4%	92.9%	0.9%	0.0%
Eisai	Pharmaceuticals & biotechnology	686	0.3%	10.6%	88.1%	1.0%	0.0%
Amazon.com	General retailers	566	94.7%	2.0%	1.2%	1.4%	0.7%
Otsuka	Pharmaceuticals & biotechnology	495	4.7%	12.2%	79.3%	3.2%	0.6%
Finmeccanica	Aerospace & defence	395	35.9%	27.5%	5.8%	28.9%	1.9%
Celgene	Pharmaceuticals & biotechnology	315	1.6%	9.1%	89.2%	0.1%	0.0%
Petroleo Brasiliero	Oil & gas producers	208	0.6%	6.1%	64.8%	14.9%	13.6%
PetroChina	Oil & gas producers	41	0.0%	3.3%	87.8%	4.5%	4.5%
Banco Santander	Banks	12	0.0%	66.7%	16.7%	0.0%	16.7%
Vale	Mining	9	11.1%	12.8%	73.3%	2.8%	0.0%

Overall, this first look at the technological competences of top Scoreboard companies suggests that further investigations on the topic would provide very interesting insights into the type of competition between them and within sectors. As we have seen there is a strong concordance, at least at a broad level, between technologies and sectors, but company specificities emerge. A more detailed analysis of technological competences would enable us to better define companies' patent portfolios, test whether this concordance holds at a more disaggregated level, and study how these technological competences are transferred.

4.2 ...and at KETs

Another interesting analysis that the matching exercise at stake allows us to carry out is to look at the specialisation of companies in technologies that are retained to have a "key" role. At the EU level, for example, the so-called *Key Enabling Technologies* (KETs) are attracting attention. These have been identified by defining a series of technologies on the basis of their potential impact at societal level ("their [superior] economic potential, contribution to solving societal challenges and knowledge intensity".[12])

In the "Feasibility study for an EU Monitoring Mechanism on Key Enabling Technologies" (2012), KETs are identified, among others, though a "diffusion approach", which basically maps 6 groups of technologies in terms of IPC code lists: 1) Nanotechnology; 2) Photonics; 3) Industrial biotechnology; 4) Advanced materials; 5) Micro- and nano-electronics; 6) Advanced Manufacturing Technologies for other KETs (pp. 25-26, ibidem).

In this study we define KETs-related patents as those with at least one IPC subclass related to a key enabling technology. Although this approach gives equal weighting to patents with different KETs "intensity", it is quite straightforward to apply and has the advantage of facilitating comparisons with the overall Patstat patents.

From the outset, let us observe that the share of KETs-related patents in our top R&D investors is 18.1%, which is very similar to that at the worldwide level (17.8%). Overall, the top 100 Scoreboard companies do not seem to be more KETs-intensive than "the rest of the world" (according to the rough definition used). Whether an R&D leadership in investments could actually be necessary/desirable to stay ahead in the development of KETs is an issue for further investigation.

Further insights emerge by looking at the KETs intensity of the sample companies by macro-sector (Figure 6). Companies operating in medium technological sectors are the most KETs-intensive, followed by those operating in Electronic and Pharma & Health sectors. On the other side of the spectrum, lower KETs intensity is shown by companies operating in the Automobile, Low tech, and Software sectors.

[12] COM(2009), 512/3, "Preparing for our future: Developing a common strategy for key enabling technologies in the EU", SEC(2009) 1257.

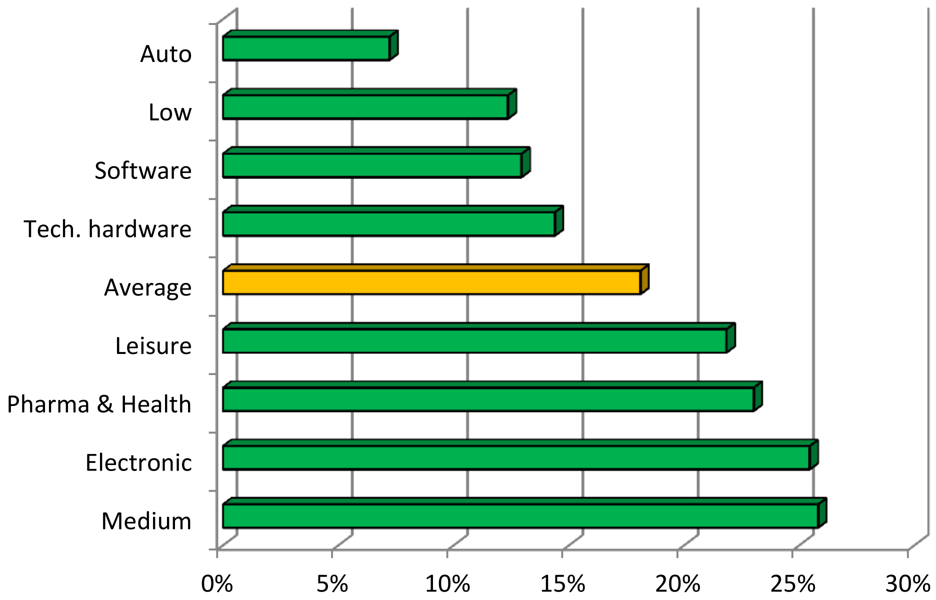


Figure 6: Share of KETs patents by sector groups

This is an interesting result, which suggests that even across top R&D spenders, sector specificities in the development of KETs matter, possibly because KETs are technologies (knowledge) develop ed and used to produce further technology (knowledge). Also in this last respect, further analysis would offer extremely relevant insights concerning Scoreboard companies´ involvement in the development of KETs.

Table 3 shows the relative performance of top 100 R&D investors based on the ranking positions on: total patent applications, R&D expenditures, total KETs-related patents and KETs intensity.

Table 3: Top 100 Scoreboard companies, ranks and KETs patents intensity

Company	Country	Patent Rank	R&D Rank	KETs Intensity	KETs Rank	KETs Intensity Rank	Rank Difference (Total - KETs)
IBM	USA	1	23	16%	3	51	-2
Samsung Electronics	South Korea	2	5	24%	1	30	1
Panasonic	Japan	3	14	26%	2	25	1
Siemens	Germany	4	19	15%	13	54	-9
General Electric	USA	5	30	17%	8	48	-3
Canon	Japan	6	37	18%	10	45	-4
Fujitsu	Japan	7	49	19%	11	43	-4
Sony	Japan	8	18	17%	14	50	-6
Toshiba	Japan	9	33	26%	4	22	5
Philips	The Netherlands	10	63	21%	9	36	1
Hitachi	Japan	11	24	25%	5	26	6
Microsoft	USA	12	2	1%	72	97	-60
FUJIFILM	Japan	13	64	28%	6	21	7
Hewlett-Packard	USA	14	47	13%	24	56	-10
Robert Bosch	Germany	15	21	8%	31	70	-16
Toyota Motor	Japan	16	1	9%	32	68	-16

Nokia	Finland	17	15	3%	57	90	-40
Intel	USA	18	8	23%	16	31	2
NEC	Japan	19	66	23%	18	32	1
Mitsubishi Electric	Japan	20	74	22%	19	33	1
BASF	Germany	21	70	41%	7	9	14
Sharp	Japan	22	75	36%	12	14	10
Google	USA	23	26	9%	37	67	-14
Alcatel-Lucent	France	24	48	11%	33	59	-9
Ericsson	Sweden	25	29	4%	54	84	-29
Honda Motor	Japan	26	13	6%	46	76	-20
Honeywell	USA	27	80	21%	26	38	1
Denso	Japan	28	40	13%	36	57	-8
Texas Instruments	USA	29	83	34%	17	15	12
General Motors	USA	30	9	4%	56	80	-26
Oracle	USA	31	31	4%	61	82	-30
Qualcomm	USA	32	50	6%	52	77	-20
DuPont	USA	33	78	46%	15	7	18
Ricoh	Japan	34	93	10%	41	62	-7
Bayer	Germany	35	38	33%	23	16	12
Huawei	China	36	41	1%	89	96	-53
United Technologies	USA	37	71	9%	45	66	-8
Cisco Systems	USA	38	22	2%	76	92	-38
Procter & Gamble	USA	39	73	9%	47	65	-8
STMicroelectronics	The Netherlands	40	67	29%	30	20	10
Pfizer	USA	41	6	21%	35	39	6
Research In Motion	Canada	42	92	2%	82	95	-40
ZTE	China	43	98	0%	95	100	-52
Boeing	USA	44	43	18%	39	44	5
Delphi	UK	45	100	11%	50	60	-5
Ford Motor	USA	46	25	3%	78	89	-32
Medtronic	USA	47	95	4%	69	83	-22
Advanced Micro Devices	USA	48	99	52%	20	5	28
Volkswagen	Germany	49	3	4%	70	85	-21
Roche	Switzerland	50	7	43%	25	8	25
LG	South Korea	51	36	41%	28	10	23
Renesas	Japan	52	61	54%	21	3	31
Continental	Germany	53	65	10%	53	61	0
Dow Chemical	USA	54	85	56%	22	2	32
GlaxoSmithKline	UK	55	17	24%	38	29	17
Broadcom	USA	56	76	7%	64	74	-8
EADS	The Netherlands	57	34	10%	58	64	-1
Nissan Motor	Japan	58	20	8%	62	72	-4
Novartis	Switzerland	59	4	25%	39	28	20
Merck US	USA	60	10	20%	42	40	18
NTT	Japan	61	44	10%	60	63	1
Sumitomo Chemical	Japan	62	89	54%	27	4	35
Johnson & Johnson	USA	63	11	15%	51	55	12
Abbott Laboratories	USA	64	35	20%	43	41	21
Aisin Seiki	Japan	65	91	7%	67	75	-2
SAP	Germany	66	58	0%	97	99	-31
Apple	USA	67	59	5%	77	79	-10
Mitsubishi Chemical	Japan	68	82	57%	29	1	39
Caterpillar	USA	69	62	3%	81	88	-12
Merck DE	Germany	70	77	40%	34	11	36
Peugeot (PSA)	France	71	45	3%	88	91	-17
Daimler	Germany	72	12	4%	82	86	-10
Sanofi-Aventis	France	73	16	22%	48	35	25

AstraZeneca	UK	74	28	15%	59	53	15
Renault	France	75	53	2%	91	93	-16
Nestle	Switzerland	76	87	7%	73	73	3
BMW	Germany	77	32	4%	84	81	-7
Volvo	Sweden	78	56	3%	87	87	-9
Bristol-Myers Squibb	USA	79	39	31%	44	18	35
Hyundai Motor	South Korea	80	81	6%	79	78	1
Monsanto	USA	81	94	26%	49	23	32
EMC	USA	82	54	2%	94	94	-12
Boehringer Ingelheim	Germany	83	46	16%	65	52	18
Eli Lilly	USA	84	27	20%	68	42	16
Royal Bank of Scotland	UK	85	86	8%	86	71	-1
Fiat	Italy	86	52	9%	85	69	1
Daiichi Sankyo	Japan	87	60	11%	80	58	7
Takeda Pharmaceutical	Japan	88	42	32%	63	17	25
Novo Nordisk	Denmark	89	90	26%	66	24	23
Amgen	USA	90	51	52%	55	6	35
Astellas Pharma	Japan	91	55	25%	74	27	17
Eisai	Japan	92	88	30%	71	19	21
Amazon.com	USA	93	68	1%	98	98	-5
Otsuka	Japan	94	72	37%	75	12	19
Finmeccanica	Italy	95	57	17%	90	46	5
Celgene	USA	96	97	17%	92	47	4
Petroleo Brasiliero	Brazil	97	96	21%	93	37	4
PetroChina	China	98	69	37%	96	13	2
Banco Santander	Spain	99	79	17%	99	49	0
Vale	Brazil	100	84	22%	99	34	1

It is interesting to notice that the ranking correlation between the R&D and patent scores of the 100 companies is not particularly strong (0.371, *see table 2*) hinting that company specificities play an important role in the R&D-patent relationship. Particularly interesting would be the investigation of these specificities in the framework of the technological regimes[13] in which companies operate, i.e. how *technological opportunities, appropriability of innovations, cumulativeness of technological advantages* and *properties of the knowledge base* (p. 391, ibidem) interact with companies' specific competences in shaping their technological evolution and economic performances.

Table 4: Spearman correlations across different rankings

	R&D Rank	Patent Rank	KET Patents Rank	KET Intensity Rank
R&D Rank	1			
Patent Rank	0.371*	1		
KET Patents Rank	0.231*	0.757*	1	
KET Intensity Rank	-0.160*	-0.013	0.580*	1

Note: ** significant at 5%*

[13] Breschi, Stefano, Franco Malerba, and Luigi Orsenigo. "Technological regimes and Schumpeterian patterns of innovation." *The Economic Journal* 110.463 (2000): 388-410.

Similarly, it is interesting to investigate whether companies with a higher number of patent applications also have a higher intensity of KETs, which would suggest that the development of the latter inventions might benefit from a wider and possibly more diversified set of technological competences. A glance at the differences in the relative rankings does not support this hypothesis, and the same holds true for the relative correlation index which is not statistically significant (-0.013). The ranking correlation between R&D and KETs patents is even lower than those between R&D and patents (0.2282); what is more, higher levels of R&D expenditure seems to be negatively associated (-0.160) with the intensity in KETs-related patenting activity. This preliminary finding is interesting and merits further investigation in order to be confirmed. In particular, the fact that the development of key enabling technologies goes along with a deeper and possibly more specialised set of technological competences, held by firms with a strong orientation towards R&D investments (R&D intensity) and active in a specific branch of the economy, seems to be an interesting research hypothesis.

5 Conclusions

In concluding the note, it is helpful to sum up the main points emerging from this first explorative analysis of the patenting activity of the top 100 Scoreboard R&D investors. With respect to each of them, further analysis can be envisaged towards a better understanding of their technological competences and of their technological portfolio.

First of all, in aggregated terms, the R&D investments and the patent applications of the investigated sample of companies both increased over the period 2004-2010. However, the share of world patents of these companies appears much smaller than that of their R&D. Furthermore, the growth of their overall patent portfolios is less than proportional of those of their R&D activities. Finally, the technological regimes in which firms operate seem to play an important role together with firms' specificities; their role in shaping the R&D-Patent relationship could be refined taking the preliminary results of this note as a starting point (e.g. the very low Patent to R&D ratio observed for companies operating in the Pharma & Health sector). In all these respects, the kind of study under exploration is worthwhile pursuing further on the basis of firm level data.

The analysis of the technological competences of the sample also yields promising results. A first inspection of the IPC classes of the patents filed by the top 100 Scoreboard sample seems to suggest a certain concentration in the kind of technological knowledge that they master. This appears to be suggested also by the most disaggregated levels of the IPC classification system and encourages us to re-examine, at disaggregated level, the “old-story” (à la Patel and Pavitt) of the little variety of the largest world investors’ competencies: do these firms really “know more than what they make?”[14] This is a second research issue that the kind of study under investigation would enable us to pursue.

Related to the previous point is the analysis of the companies’ involvement in specific technological fields, as for example in the so-called Key Enabling Technologies. Matching company and patent data looks promising also in this last respect. For example, sector specificities, probably deriving from the interaction of the knowledge base required to develop particular technologies (e.g. related to Advanced manufacturing) and that already mastered by the companies operating in a given sector, seem to emerge. Further research could help exploring this issue by incorporating firm-specific considerations in the analysis.

Augmenting the EU R&D Investment Scoreboard with patent data actually proves to be an interesting extension to be pursued. Research questions to be addressed include: R&D leaders (followers) appear to stay ahead (lag behind) in terms of patent applications as well: would this remain the case for the full Scoreboard? More generally, how do country and sector specificities operate in a patent-augmented version of the IRI Scoreboard?

Moreover, a relevant extension of this preliminary analysis is to measure the “quality” of companies' patents, which has proved to vary widely from patent to patent and both at firm and industry levels (Scherer, 1965). Therefore, weighting patents by some quality measure would allow for more straightforward company and sector comparisons. These measures, together with the financial data of the Scoreboard companies, will enable us to assess the economic value and technological potential of the companies' patent portfolios, that is, the impact that these (might) have on firms' economic performances and on further technological developments.

To summarise, while it has helped in facing and solving (or trying to solve) the several delicate methodological issues posed by the underlying matching (in particular those summarised in Technical Annex I), this pilot study confirms the feasibility of the exercise and the interest in extending it to the full sample of Scoreboard companies.

[14] Stefano Brusoni, Andrea Prencipe and Keith Pavitt (2001).

References

Breschi S., Malerba F. and Orsenigo L. (2000), "Technological regimes and Schumpeterian patterns of innovation", *The Economic Journal,* 110.463, 388-410.

Breschi S., Lissoni F., & Malerba F. (2003). "Knowledge-relatedness in firm technological diversification", *Research Policy*, 32(1), 69-87.

Brusoni S., Prencipe A. and Pavitt K. (2001), “Knowledge specialization, organizational coupling, and the boundaries of the firm: Why do firms know more than they make?”, *Administrative Science Quarterly*, 46(4), 597-621.

EC (2012), "Feasibility study for an EU Monitoring Mechanism on Key Enabling Technologies".

Frietsch R., Neuhäusler P., and Rothengatter O. (2012), "*Patent Applications – Structures, Trends and Recent Developments",* Studien zum deutschen Innovationssystem Nr. 8-2012 Fraunhofer Institute for Systems and Innovation Research ISI, http://www.e-fi.de/fileadmin/Innovationsstudien_2012/StuDIS_08_ISI.pdf

Magerman T, Grouwels J., Song X. and Van Looy B. (2009). "*Data Production Methods for Harmonized Patent Indicators: Patentee Name Harmonization",* EUROSTAT Working Paper and Studies, Luxembourg

Patel P. and Pavitt K. (1997), “Technological competencies of the world's largest firms: Complex and path-dependent, but not much variety”, *Research Policy*, Volume 26, Issue 2, 141–156.

COM(2009), 512/3, “Preparing for our future: Developing a common strategy for key enabling technologies in the EU”, SEC(2009) 1257.

Scherer, F. M. (1965). “Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions”, *American Economic Review*, 55(5): 1097-1125.

Schmoch U. (2008), “Concept of a Technology Classification for Country Comparisons”, Geneva, World Intellectual Property Organisation (WIPO).

Thoma G., Torrisi S., Gambardella A., Guellec D., Hall B.H. and Harhoff D. (2010). *Harmonizing and combining large datasets – An application to firm-level patent and accounting data.* NBER working paper series 15851; http://www.nber.org/papers/w15851

Wilson D (2003), “Are We Running out of New Ideas? A Look at Patents and R&D”, *FRBSF Economic Letter*.

Technical Annex I

Data sources involved

This annex aims to describe how patents have been assigned to the top 100 R&D world investors reported by the IRI Scoreboard.

The primary source of information about patents is PATSTAT, a database produced by the European Patent Office (EPO) OECD and Eurostat. In the present exercise the match has been extended to all application authorities although extensive usage has been restricted to three of them (EPO, USPTO and PCT). Nevertheless the methodology applied might be extended to a wider pool of companies. Other matching exercises had been previously been performed and our work widely benefited from such efforts. For instance in Grid Thoma et al (2010) a similar methodology is described where the match is applied to Amadeus and, on the Patstat side, is restricted to patents applied for at European Patent Office.

Datasets harmonisation

The quality of data sources was extremely different: while Scoreboard data were normalised, PATSTAT applicant names are very raw. In the latter database company names may appear with different spellings: the typical example is IBM that may appear also as International Business Machines, I.B.M. or Int Busn Mach, making difficult to regroup under a single label patents applied for by the same entity.

In order to overcome this difficulty we decided to use the ECOOM-EUROSTAT-EPO PATSTAT Person Augmented Table (EEE-PPAT), which relies on a comprehensive method to arrive at harmonised patentee names in an automated way (described in detail in Magerman et al. (2009)). The EEE-PPAT table adds an important piece of information: the assignee sector allocation, which enables us to identify whether patentees are private business enterprises, universities/higher education institutions, governmental agencies or individuals. In such a way applicants flagged as individual have been excluded from our exercise.

Both data sources have been completely processed in order to remove nonstandard ASCII characters, double spaces, other common misspellings/typographical errors and to remove the legal designation (LLC, INC, CORP…) shifting it into a different field. The same processing algorithm has been applied to both datasets in order to assure that the results are comparable (name are standardised using the same procedure).

Matching the data

Due to the high percentage of standard names, and the fact that for company names a small name change may occur among distinct companies (since most of them are acronyms), it has been decided not to use any edit distance criteria (like Levensthein or N-gram functions) but to rely only on three criteria:

1) Perfect match: where names, removing legal designation, are exactly the same

2) Alphanumeric match: where the names, keeping only [A-Z] and [0-9] are the same (FI: I.B.M. = IBM = I B M)

3) Jaro-Winkler distance: names are broken into tokens and the similarity score is computed by the number of tokens in common, weighted on the inverse of frequency. The higher the Jaro–Winkler distance for two strings, the more similar the strings. Only results above a threshold value have been considered valid matches.

Eventually, only the best result could be considered as a match, since it is very likely that the databases could contain both the main company and a subsidiary, where the match among the same entity would turn out to be type 1 or 2 (perfect or alphanumeric) while non-homogenous entities would match with type 3 (tokens). This criterion avoids for instance Harley Davidson Motor Company and Harley Davidson Motor Company Financial Services to be matched.

Aside from name, country codes have also been taken into account as main criteria for match, because the same applicant name in different nations may refer to distinct entities (FI Ministry of health). Since in PATSTAT, apart from EPO, coverage of country code is sometimes very poor, this data is also enriched using priorities and family data (where in patent family data homonymous with a non-blank country code exists, such data is retained).

Note that, since the time frame of two datasets is different (Scoreboard contains latest data available, PATSTAT names are taken at the moment of application or grant of patent), previous names in Scoreboard and 'Also Known As' names have also been included in the match, for companies where this piece of information is available.

Filtering the match

The Filtering phase of the algorithm is meant to discriminate false matches and fine-tune the parameters (e.g. the threshold value for tokens similarity).

Whenever trying to match two distinct data sources an inevitable trade-off arises between flagging positive matches as false negatives and including false positives as true matches. In the first case an existing link between a patenting and a financial entity would not be established, in the second case a non-existing link would be established. In computer science this is measured by two indicators, the precision and recall rates, which are defined as:

Recall rate = true positives / (true positives + false negatives)

Precision rate = true positives / (true positives + false positives)

Ideally both values should be as near as possible to one, but in reality these rates are used to balance filtering criteria in order to allow the most satisfactory solution among those available. For this purpose a hand check has been carried out on a control sample of the total scoreboard population involved in the match (about 3,500 company names) seeking them on ESPACENET (online patent database provided by EPO), also cross-checking other patent data or company websites in order to have an accurate specimen to be used as a benchmark for the overall algorithm. Results on cross-checking have been used for balancing the algorithm.

Eventually other criteria were used in order to highlight possible false positives. Such criteria relied on:

a) Same technology fields (cross-check International Patent Classification of applications assigned against NACE sector of the Scoreboard company)

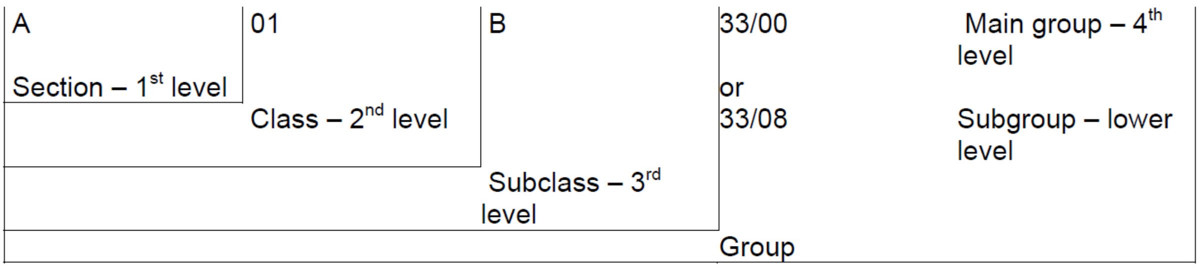
b) Time frame (check application year with start-up year of the Scoreboard company)

c) Other scoreboard relevant data (R&D expenses > 0)

Technical Annex II

The IPC classification has a hierarchical structure (different levels). In the following figure it is illustrated how different parts of a code correspond to different levels of the classification.

For a guide to IPC classification click here.



For instance the code depicted in the figure A01B33/00 (stored in the database as A01B 33/00) should be read as:

A — Human Necessities

01 - Agriculture; Forestry; Animal Husbandry; Hunting; Trapping; Fishing

B - Soil Working in Agriculture or Forestry; Parts, Details, or Accessories of Agricultural Machines or Implements in General

33/00 - Tilling implements with rotary driven tools

(33/08 - Tools; Details, e.g. adaptations of transmissions or gearings)

Technical Annex III

"The concept of the patent family first emerged through the Paris Convention on the Protection of Industrial Property in 1883, while automated systems enabling patent family searching became available through the establishment of the IIB in The Hague in 1947 and INPADOC in Vienna in 1972. [...] An important point when using any database to retrieve information on patent families is that there is never any guarantee that you will find all the corresponding patent documents that exist (EPO website)". For this study we have opted for the INPADOC family definition because it allows us to match more than 98% of the patent applications of our sample.

The extended (INPADOC) patent family groups together all documents directly or indirectly linked via a priority document. The case below illustrates how a patent family is built up starting from different documents. In this example documents D1-D5 belong to the same patent family, P1.

Document (patent)		Priorities	
Document D1	Priority P1		
Document D2	Priority P1	Priority P2	
Document D3	Priority P1	Priority P2	
Document D4		Priority P2	Priority P3
Document D5			Priority P3

In the present exercise, we had 1,582,296 different patent applications over the period 2000-2010 that, once converted into INPADOC families, become 906,073 "unique patents (or invention) observations". This means that on average each family contains about 1.75 patent applications, although it should be noted that sector specificities arise. In particular, firms operating in the Pharma & Health sector have a higher average family size, followed by those operating in the Low technological sectors, and show a particularly high standard deviation. This is mainly due to some outliers in the family size distribution, in fact 3 out of the 5 biggest families belong to firms operating in the Pharma & Health sector: Roche (biggest family size: 2,512), GlaxoSmith (448) Merck US (348). These outliers alone could be an interesting objective of study.

Table AIII: Average size (and standard deviation) of families by sector group

Sector Group	Average	Std. dev.
Pharma & Health	2.5	10.5
Low	2.0	1.7
Leisure	1.9	1.5
Medium	1.8	1.9
Tech. hardware	1.7	1.6
Auto	1.5	0.9
Software	1.5	1.1
Electronic	1.5	1.0

European Commission
EUR 26350 – Joint Research Centre – Institute for Prospective Technological Studies

Title: The patenting activity of the top IRI Scoreboard Companies: an introductory note

Authors: Antonio Vezzani, Fabio Montobbio, Sandro Montresor and Gianluca Tarasconi

Luxembourg: Publications Office of the European Union

2014 – 24 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424 (online)

ISBN 978-92-79-34744-3 (pdf)

doi:10.2791/53758

Abstract

The present note contains an explorative, introductory analysis of the patenting activity exhibited by the top 100 companies of the IRI Scoreboard, and intends to identify strengths and weaknesses for its possible future extension to the whole Scoreboard. With respect to these companies, patent data are drawn from Patstat, on the basis of which patent families are built up, and crossed with other data on their R&D investments. Both the R&D and the patent applications of the investigated sample of companies increase over time. At the same time, important sector specificities in the R&D-patent relationship have been found. The analysis of the technological competences of the overall sample yields promising results. A first examination of the IPC classes of the patent applications suggests a certain concentration in the kind of technological knowledge that companies master. The analysis of the knowledge base and, more specifically, the companies' involvement in the creation of key enabling technologies (KETs) also highlights that important sector specificities go along with firm-specific factors. All in all, "augmenting" the Scoreboard data with company level patent information appears to be an interesting extension to be pursued.

LF-NA-26350-EN-N

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

Key policy areas include: environment and climate change; energy and transport; agriculture and food security; health and consumer protection; information society and digital agenda; safety and security including nuclear; all supported through a cross-cutting and multi-disciplinary approach.

Publications Office

ISBN 978-92-79-34744-3