Research Paper

Gauging a Firm's Innovative Performance Using an Integrated Structural Index for Patents

Xiaojun Hu[1] & Ronald Rousseau[2,3†]

[1]Medical Information Center, Zhejiang University School of Medicine, Hangzhou 310058, China
[2]Department of Mathematics, KU Leuven, Leuven B-3000, Belgium
[3]Information and Library Science (IBW), University of Antwerp (UA), Antwerp B-2000, Belgium

Citation: Xiaojun Hu & Ronald Rousseau (2016). Gauging a Firm's Innovative Performance Using an Integrated Structural Index for Patents.
Received: Jan. 28, 2016
Revised: Feb. 22, 2016
Accepted: Feb. 25, 2016

Abstract

Purpose: In this contribution we try to find new indicators to measure characteristics of a firm's patents and their influence on a company's profits.

Design/methodology/approach: We realize that patent evaluation and influence on a company's profits is a complicated issue requiring different perspectives. For this reason we design two types of structural h-indices, derived from the International Patent Classification (IPC). In a case study we apply not only basic statistics but also a nested case-control methodology.

Findings: The resulting indicator values based on a large dataset (19,080 patents in total) from the pharmaceutical industry show that the new structural indices are significantly correlated with a firm's profits.

Research limitations: The new structural index and the synthetic structural index have just been applied in one case study in the pharmaceutical industry.

Practical implications: Our study suggests useful implications for patentometric studies and leads to suggestions for different sized firms to include a healthy research and development (R&D) policy management. The structural h-index can be used to gauge the profits resulting from the innovative performance of a firm's patent portfolio.

Originality/value: Traditionally, the breadth and depth of patents of a firm and their citations are considered separately. This approach, however, does not provide an integrated insight in the major characteristics of a firm's patents. The $S_h(Y)$ index, proposed in our investigation, can reflect a firm's innovation activities, its technological breadth, and its influence in an integrated way.

Keywords Patent analysis; Structural h-index; Market value of patents; Technological value of patents; Pharmaceutical industry; Nested case-control

JDIS
Journal of Data and Information Science
Vol. 1 No. 1, 2016
pp 6–27
DOI: 10.20309/jdis.201603

† Corresponding author: Ronald Rousseau (E-mail: ronald.rousseau@kuleuven.be).

1 Introduction

The technological scope of a firm's patents, as expressed by the number and nature of the classes to which these patents are assigned, is an important element to describe the relation between a company's technological diversity and its profits (Chen, Jang, & Wen, 2010; Chiu, et al., 2010; Olivo et al., 2011). Indeed, research suggests that the scope of patents owned by a firm has a strong impact on performance and is, as such, an economically significant variable (Lerner, 1994; Reitzig, 2003).

As we want to take an international point of view we use the International Patent Classification (IPC) codes, but not American or European patent codes. Moreover, IPC codes have already been used in several other investigations (Chen, Jang, & Wen, 2010; Chiu et al., 2010; Lerner, 1994; Sapsalis, van Pottelsberghe de la Potterie, & Navon, 2006). Following these colleagues we use the number of 3- or 4-digit IPC codes assigned to a patent as a proxy of its technological breadth. Besides, the depth of a patent is also a structural element involved in a patent portfolio. Consider, for example, an IPC code such as "A61K-037": the head 3 to 4 digits refer to a technological class and subclass (A61K), and the tail digits reflect the technological depth of the patent involved (037). This suggests that at the structural level, the breadth of patent is the primary structure, and the depth of a patent is the secondary one.

The ratio between the total number of codes (7- or 8-digit codes) used to describe patent p and the number of classes and subclasses, reflected by 3- or 4-digit codes, is called its technological depth, denoted as $d(p)$. It is at least one and usually strictly larger than one. This indicator is not very precise because the number of 7- and 8-digit IPC codes is quite different per class (Lodh & Battaggion, 2014; Zhang, Chen, & Niu, 2012).

Generally, the broader the scope of a patent, the larger the number of competing products and processes that might infringe on the patent (Merges & Nelson, 1990). In this context, these authors pointed out that excessively broad patents may lead to use by other parties. Yet, Gilbert and Shapiro (1990) claimed that broader patents provide inventors with a greater ability to earn profits. As the competitive strength of a firm's patents is an aspect of their market value, technological value, and social value, finding the optimal depth and breadth of a patent is a complex as well as a controversial topic (Guan & Gao, 2009; Hu & Rousseau, 2015; Hu, Rousseau, & Chen, 2012; Klemperer, 1990; Lee, 2009; Palokangas, 2011; Reitzig, 2003). We recall that, according to Gilbert and Shapiro (1990), the breadth of a patent is related to the flow of profits available to the patentee as well as to the minimum improvements that another inventor has to make in order to obtain a non-infringing patent. According to Lerner (1994) the market value of patents, sometimes even of a single patent, can have a major effect on the value of a firm. Exploring the optimal depth and breadth of a patent, researchers have increasingly recognized the importance to focus on the breadth of a patent (Denocolò, 1996; Kanniainen & Stenbacka, 2000; Merges & Nelson, 1990; Palokangas, 2011).

Continuing our research on the characteristics of the *IPCh* indicator (Hu & Rousseau, 2015) (its definition is recalled further on), the purpose of this contribution is:

1) To show, using a large dataset, how the *IPCh* indicator for patents is able to provide information on a company's innovative activities;
2) To provide convincing evidence that the *IPCh* and the yearly h-index of patents are closely related to a firm's innovative performance, and compare this with a synthetic indicator including the depth of a patent, based on companies in the pharmaceutical industry; and
3) To provide a simple way to gauge a firm's patent performance by jointly taking two h-type indices into account, each reflecting another aspect of the h-core in the lists of technological breadth and citations (reflecting market value and technological value).

As we are aware of the shortcomings of all h-type indices (Bouyssou & Marchant, 2011; Waltman & van Eck, 2012), we nevertheless claim that our approach is a useful addition to the patent toolbox. Moreover, no indicator on its own can provide information from all possible perspectives at the same time. Borrowing the terminology of Valiant (2013), proposed by him in the context of machine learning, the information provided by such an indicator is at best Probably Approximately Correct (PAC).

2 A Short Literature Review Related to the Concepts Used in This Contribution

2.1 The General h-index Idea

Hirsch (2005) proposed the h-index as an author-level indicator combining productivity (published articles) and impact (received citations). Soon his idea was applied to other source-items relations such as journal publications and citations (Braun, Glänzel, & Schubert, 2005), a company's patent assignments and their citations in other patents (Guan & Gao, 2009), publications and citations of topics, restricted to recent years (Banks, 2006) or availability of books and their loans according to a library classification (Liu & Rousseau, 2009). We first recall the basic mechanism for calculating the h-index of an actor (author, company, or a journal). One considers a two-dimensional table of sources and items, where sources, e.g.

publications or patents, are ranked according to items, e.g. received citations. Sources with the same number of items are given different rankings, but the exact order does not matter. Then actor A's h-index is equal to the number h if the first h sources have each at least h items, while the source ranked h+1 has strictly less than h+1 items.

2.2 Patent Analysis

The relation between the breadth and depth of its patents on the one hand, and the health of a firm on the other, has been studied for several decades (Denicolò, 1996; O'Donoghue, Scotchmer, & Thisse, 1998; Palokangas, 2011; Prencipe, 2000; Wang & von Tunzelmann, 2000). Yet, no final answer about the optimal breadth and depth of patents has been found (Ozman, 2007; Zhang, Chen, & Niu, 2012; Lodh & Battaggion, 2014; Breschi, Lissoni, & Malerba, 2003). When using diversity indexes to measure the technological breadth and depth of a firm, it may happen that results are biased downwards for small and medium-sized firms for which the scale of technological activities is small (Chen, Jang, & Wen, 2010; Hu & Rousseau, 2015; Miller, 2006; Palokangas, 2011). Moreover, diversity indices such as the Rao-Stirling index may show cyclical patterns that are not related to a company's profits but are rather related to the number of inventors (Leydesdorff, 2015). This suggests that if one wants to understand the optimal breadth and depth of patents, an approach different from the "complexity and diversity" might be worth investigating (Lodh & Battaggion, 2014; Wang & von Tunzelmann, 2000).

Traditionally, the breadth and depth of patents of a firm and their citations are considered separately. This approach, however, does not provide an integrated insight in the major characteristics of a firm's patents. It has been observed that return on investment of a patent depends largely on a firm's market value and its technological value, while the competitive strength of a firm's patents bears a close relation to market value, technological value, social value of patents, and healthy management styles (Guan & Gao, 2009; Hu & Rousseau, 2015; Lee, 2009; Palokangas, 2011).

3 Methodology

We develop a new approach to gauge a firm's innovative performance based on the following insights.

3.1 Potential Applications of Patents

We claim that one of the most important elements affecting the potential applications of a patent is its breadth, operationalized by codes, such as the IPC, the

U.S. Patent Classification System (USPC), Cooperative Patent Classification (CPC) or the European Patent Office (EPO) codes assigned to it. This set of codes forms a basic aspect to grant its owner either a very limited right to exclusive use or a more general right covering a variety of different realizations of the invention (Reitzig, 2003). This fact implies that patents can differ with respect to the degree of protection afforded to an invention (Gilbert & Shapiro, 1990; Klemperer, 1990). In this context we note that accrediting codes to a patent is an arena in which patent examiners exercise wide discretion. In general, the broader the patent, the higher the chance to be applied in different practical fields and the larger the potential profits to the firm or a purchaser of the firm's patent (Palokangas, 2011). This leads to the claim that the optimal breadth of patents should focus on a firm's performance. Excessively broad patent claims increase the patentees' non-market related risks from rivals and provide them with little flexibility to face unexpected situations (Merges & Nelson, 1990). However, the narrower a patent's claims, the more the patentee may be victim of imitation as very similar products may lie outside the original patent's claims (Denicolò, 1996; Kanniainen & Stenbacka, 2000).

A firm which focuses on excessively broad patents would overspend its research and development (R&D) capital by developing or buying an overly large number of patents. And, *vice versa*, if most of the firm's patents are of narrow breadth, the firm reduces its chance to earn larger profits than competitors. Obviously, these two extreme cases do not lead to healthy management styles in a competitive industry. Therefore, it is very important to measure the competitive strength of patents and hence the "weight" of a firm's patent portfolio. Such an investigation must include the number of patents, their impact and their breadth.

3.2 The Structure of Patents and Their Influence Must Jointly Be Taken into Account

It is well known that the received number of patent citations is an important indicator to measure the influence of a patent. Moreover, patent citations have a positive relation with the profits of the patent owner (Hu, Rousseau, & Chen, 2012; Trajtenberg, 1990).

Many investigations point out that, compared to the breadth of a patent (the primary dimension), it is less meaningful to focus on the depth of a patent because the determination of a patent's depth is just approximate and no positive relation between a patent's performance and its depth has been found (Gilbert & Shapiro, 1990; Kanniainen & Stenbacka, 2000; Klemperer, 1990; Lodh & Battaggion, 2014; Ozman, 2007; Palokangas, 2011; Reitzig, 2003; Zhang, Chen, & Niu, 2012).

Grönqvist (2009) argues that broader patents are not necessarily more valuable than narrower ones. Concretely, patents described with many codes do not necessarily lead to a larger profit for the firm. Therefore, neither the breath of patents nor the number of received citations on their own are clear-cut indicators for the value of a company's patent portfolio. If we want to understand the competitive strength of a firm from the perspective of patent performance, the primary structure of patents (patent breadth), the secondary structure (patent depth), and their influence should jointly be taken into account in a multi-layered approach (Denicolò, 1996; Hu, Rousseau, & Chen, 2012; Palokangas, 2011). Abstractly, their relationships can be described with Equation (1):

$$SP = f\,(p,bp,dp,cp), \tag{1}$$

where SP denotes the competitive strength of patent-related performance of a firm, and p is the number of patents; bp denotes their breadth, dp their depth and cp the number of received citations.

3.3 The Structural h-index for Patents

To reveal the relation between the essential structure of patents and their competitive strength, e.g. profit performance, in the real world, and clarify the controversy on the influence of depth on a patent's profit, we propose two types of structural h-indices for patents: (1) the structural h-index, a primary one, combining the number of patents with the primary structure (breadth of patent) and with forward, i.e. received, citations; (2) the synthetic structural h-index, using the number of patents, the breadth and depth of these patents, and the number of forward citations.

Hence, we hypothesize that the primary structure of patents (patent breadth) and their influence on a firm can be measured by a structural h-type index, combining different aspects in a dynamic way.

3.4 Definitions of *IPCh* and Yearly h-index

A firm's innovation activities are operationalized as the number of patents, while their technological breadth is operationalized by the number of 3- or 4-digit IPC codes. Consider a set of patents granted to a firm in a certain year Y, ranked in decreasing order of the number of 3- or 4-digit IPC codes assigned to them. Then the IPC h-index of this firm in the year Y is equal to q if q is the highest rank such that the first q patents are assigned to at least q IPC codes (Hu & Rousseau, 2015). The resulting indicator is denoted as $IPCh_3$ or $IPCh_4$ depending on the number of digits that have been used.

Next, we define a yearly h-index slightly modified from the original meaning of Hirsch (2005) to map a firm's innovation activities and influence in the year Y. The yearly h-index of a firm in the year Y, denoted as h_Y, is equal to h if h is the largest rank such that the first h patents receive each at least h citations within a given citation window. In the examples investigated below the citation windows always end on May 20, 2014.

3.5 Definition of the Patent Depth Yearly h-index (Dh_Y)

Next, we define the yearly h-index of patent depth in the year Y, denoted as Dh_Y as follows. Consider the set of patents granted to a firm in the year Y, ranked in decreasing order of their technological depth index, $d(p)$. The Dh_Y index of this firm in the year Y is equal to k if k is the highest rank such that the first k patents have at least a technological depth equal to k.

3.6 The Structural h-index for Patents

We define the structural h-index for patents granted in the year Y, denoted as $S_h(Y)$, as a combination, actually a multiplication, of the *IPCh* and the yearly patent h-index. Hence $S_h(Y)$ can be calculated with Equation (2):

$$S_h(Y) = IPCh_s \times h_Y, \tag{2}$$

where $s = 3$ or 4. Moreover, although not indicated in the notation, $S_h(Y)$ is time dependent, i.e. depends on the citation window, which in our examples ends on May 20, 2014. The $S_h(Y)$ sequence shows a firm's innovation activities and their technological breadth, as well as the influence of patents (by citations) in each year. As such we claim that it can be used to gauge the "primary weight" of a firm's patents. This claim is investigated in the next section.

3.7 The Synthetic Structural h-index for Patents

Finally, we define the synthetic structural h-index for patents granted in the year Y, denoted as $SS_h(Y)$, as a summary indicator constructed from the *IPCh*, the yearly patent h-index (h_y), and the patent depth yearly h-index and it can be calculated with Equation (3):

$$SS_h(Y) = w_1 IPCh_s + w_2 h_y + w_3 Dh_y \,, \tag{3}$$

where w_1, w_2 and w_3 are positive weights such that $w_1 + w_2 + w_3 = 1$.

4 An Application and an Empirical Study in the Pharmaceutical Industry

We recall that the pharmaceutical industry is a high-tech industry in which a firm's performance (and profit) is closely connected to the market value of its patents (Hu, Rousseau, & Chen, 2012; Chen, Shih, & Chang, 2013). Therefore, the pharmaceutical field is a good test bed to study the practical value of the new indicators $S_h(Y)$ and $SS_h(Y)$. We intend to find out if these two indicators are indeed able, as we hypothesize, to detect the "weight" of a firm's patents through their relation to a firm's profits.

4.1 Choice of Firms

The general range of firms acceptable for our purposes contains those pharmaceutical companies listed in Fortune 500 2006–2010 issued by the CNNMoney website[①]. These companies are the primary focus of our investigation, because yearly ranks for "pharmaceutical industry" are available during these years.

As there are many invisible factors affecting the performance of patents, we try to control for external variables by considering the following criteria for inclusion in our case study.

1) Firm location: Different countries have different regulations for patents which may influence realized profits (Chen, Shih, & Chang, 2013). For this reason only US companies were selected.
2) Firm internationality: Prior literature has found that there is a significant effect of firm scale on profits (Chen, Jang, & Wen, 2010). Accordingly, only US-based multinational firms included in Fortune 500 qualify.
3) Firm age: It has been shown that, in terms of innovation activities, older firms have a stronger foundation than younger ones. Hence, a firm's age influences the outcome of its patents' performance. For this reason we included only firms founded before the year 1990 (Banerjee & Cole, 2010; McMillan & Thomas, 2005).
4) Patent age: As the time between applying for a pharmaceutical patent and its return on investment is generally between 8 and 12 years, with 5 years as a strict minimum (ISTIS, 2003), and the protection period given by a patent is at most 20 years (WIPO, 2000), care must be exerted to take these facts into account (Chen, Jang, & Wen, 2010; Hu, Rousseau, & Chen, 2012). For this reason, we included only patents granted during the period 1990–2005, and considered profits reported by Fortune 500 for the period 2006–2010.

Taking all these requirements into account resulted in eight US-based multinational pharmaceutical companies meeting all the criteria, namely Johnson & Johnson, Pfizer, Merck, Bristol-Myers Squibb, Amgen, Genzyme, Allergan and Biogen Idec.

4.2 Data Collection and Processing

We extracted from the Derwent Innovations Index (DII) all patents granted to these eight companies during the period 1990 – 2005. For each record we downloaded

[①] http://money.cnn.com/magazines/fortune/fortune500/

all fields, including IPC-codes and citations received (so-called forward citations). Data were extracted on 20/05/2014. This led to a total of 19,080 patents for the eight firms. Next, we collected the yearly profits for each company as reported by Fortune 500 2006–2010.

For the dataset of a company's patents, we first counted the number of 4-digit IPC codes for each record via a simple program written by ourselves, and determined the yearly *IPCh* and yearly h-index during 1990–2005 for each company (Appendix Tables A1–A3). Then, we calculated the yearly $S_h(Y)$ and yearly $SS_h(Y)$ for each firm according to Equations (2) and (3). As the breadth of a patent is a primary structure while its depth is a secondary one, and because research suggests that both breadth and number of citations have positive relations with the profits of the patent owner, we take all these factors into account. Moreover, as previous research pointed out that 4-digit codes and citation-weighted counts can be taken as "patent-equivalents" (Miller, 2006), we – tentatively – weighted them higher than Dh_Y according to a weight of 0.4 for *IPCh* and for h_Y, and a weight of 0.2 for Dh_y in Equation (3) (Appendix Table A4).

To compare results based on 3-digit IPC codes with those based on 4-digit codes, we also collected the number of 3-digit codes for each patent (Appendix Tables A2 and A3), and calculated the corresponding S_h index.

4.3 Statistical Methods

To observe the relationship between the $S_h(Y)$ and a firm's profits, we use two different statistical methods:

1) We calculated the Spearman rank correlation coefficient between the eight companies, mean $S_h(Y)$ and mean $SS_h(Y)$ values and total profits over the period from 2006 to 2010.
2) A nested case-control (NCC) study. This type of study is an observational study whereby a case-control approach is employed within an established cohort (Bornehag et al., 2004). This is a popular and valid approach in medical studies for small-sample investigations. As such we consider it also suitable to our study. The nested case control model as applied in medical investigations is less expensive, but less efficient than a full-cohort analysis. However, it has been shown that with four controls per case and/or stratified sampling of controls, relatively little efficiency may be lost (Goldstein & Zhang, 2009).

To apply the NCC method, the eight companies are grouped according to their profits: Group H (high profit) consists of the four companies with the highest profit;

Group L (low profit) consists of the four companies with the lowest profits. For each group, we re-rank companies by their profits in a descending way and denote them GHR1, GHR2, GHR3, GHR4, GLR1, GLR2, GLR3, and GLR4 (Table 1). In this way, case-control is performed between four control-pairs of companies with the same rank order in the respective groups (such as GHR1 *vs* GLR1), and the nested control is designed by a sequence of time points, that is, yearly S_h and yearly SS_h among controlled cases between two groups during the period 1990–2005. Hence, 16 time points in total are used as observations. We recall that the $S_h(Y)$ indicator is time dependent. For example, in our case, the $S_h(Y)$ of the year 1990 has a citation window from the year 1990 to May 20, 2014, and the $S_h(Y)$ in the year 1991 has a citation window from the year 1991 to May 20, 2014, and so on. As pointed out above, such a stratified sampling of controls can lead to an efficient result.

Table 1. Controlled cases design for companies included in NCC study.

Group H				Group L			
Company	Code	Profits	Rank	Company	Code	Profits	Rank
Johnson & Johnson	GHR1	11,451.00	1	Amgen	GLR1	3,718.20	1
Pfizer	GHR2	10,461.00	2	Biogen Idec	GLR2	554.10	2
Merck	GHR3	6,610.02	3	Allergan	GLR3	395.24	3
Bristol-Myers Squibb	GHR4	4,521.80	4	Genzyme	GLR4	349.72	4

Note. Profits 2006–2010 in millions of US dollars (average per year).

Then, we compare the yearly S_h and yearly SS_h for each company during the period 1990–2005 between two groups using a Paired Samples Test, where pairs consist of a company from GH and a corresponding company from GL, as a so-called 'control.'

4.4 Results

In this section, we present the results obtained from our analysis of the 19,080 patents. We will show that the two types of structural h-indices $S_h(Y)$ and $SS_h(Y)$ have significant correlations with a firm's profits as given by Fortune 500 2006–2010. Moreover, the $S_h(Y)$ index has more significance than $SS_h(Y)$.

4.4.1 Yearly Values of S_h for Eight Companies during 1990–2005

Tables 2 and 3 show the resulting yearly S_h values. We would like to point out that the rank order of these eight companies is different from those obtained from the *IPCh* and from the h-indices separately (Appendix Tables A1–A3). We consider S_h to represent the primary competitive strength of a firm's patents.

Table 2. Yearly S_h indices of eight companies during the period 1990–2005 (using $IPCh_4$).

Year	Johnson & Johnson	Pfizer	Merck	Bristol-Myers Squibb	Amgen	Biogen Idec	Allergan	Genzyme
1990	192	161	150	85	32	1	44	12
1991	174	144	186	138	72	0	100	12
1992	145	162	224	100	40	1	100	20
1993	224	132	256	125	35	1	76	12
1994	217	150	240	182	78	30	85	35
1995	240	108	240	162	84	4	92	35
1996	328	120	280	156	84	4	84	84
1997	312	140	264	138	128	30	95	70
1998	280	174	280	174	120	66	70	90
1999	240	196	272	192	78	35	56	96
2000	203	240	200	203	98	70	64	90
2001	189	208	208	186	120	60	130	78
2002	273	252	240	208	140	91	115	60
2003	234	175	333	189	136	56	108	60
2004	210	132	296	189	105	66	90	66
2005	288	108	270	114	91	56	102	30
Mean	234.31	162.63	246.19	158.81	90.06	35.69	88.19	53.13
Rank	2	3	1	4	5	8	6	7

Table 3. Yearly S_h indices of eight companies during the period 1990–2005 (using $IPCh_3$).

Year	Johnson & Johnson	Pfizer	Merck	Bristol-Myers Squibb	Amgen	Biogen Idec	Allergan	Genzyme
1990	160	92	100	68	24	1	33	8
1991	145	120	155	92	36	0	80	12
1992	116	108	160	80	30	1	80	15
1993	160	88	192	100	21	1	57	8
1994	155	100	180	104	39	18	51	21
1995	150	54	150	108	56	4	69	21
1996	246	96	175	104	56	4	63	56
1997	234	80	198	92	64	24	57	56
1998	200	116	175	116	75	44	42	60
1999	160	112	204	128	52	21	42	64
2000	145	150	150	116	70	56	48	60
2001	135	104	130	124	60	40	78	52
2002	195	140	180	130	70	52	69	40
2003	156	100	222	135	85	40	72	40
2004	150	88	222	108	60	44	72	44
2005	216	72	210	95	65	40	68	15
Mean	170.19	101.25	175.19	106.25	53.94	24.375	61.31	35.75
Rank	2	4	1	3	6	8	5	7

4.4.2 Yearly Values of SS_h for Eight Companies during 1990–2005

Table 4 shows the yearly values of the synthetic structural h-indices for eight companies. Note that $SS_h(Y)$ combines the *IPCh*, the yearly patent h-index, and the yearly h-index of patent depth. Therefore, it reflects the first ranked patents in three essential dimensions. We may say that $SS_h(Y)$ represents the essential competitive strength of a firm's patents. It turns out that the ranks of the mean $SS_h(Y)$ for eight companies are very similar to those according to $S_h(Y)$. Only the first and the second company change places.

Table 4. Yearly SS_h indices of eight companies during the period 1990–2005 (using $IPCh_4$).

Year	Johnson & Johnson	Pfizer	Merck	Bristol-Myers Squibb	Amgen	Biogen Idec	Allergan	Genzyme
1990	15.8	11.396	13.4	9.4	5.4	1.0	6.4	3.2
1991	14.8	11.812	15.8	12.4	7.8	0.0	10.8	3.2
1992	14.4	12.740	16.8	10.6	6.4	1.0	10.6	4.2
1993	16.2	10.836	17.2	12.8	5.6	1.0	10.0	3.2
1994	15.8	12.188	17.2	14.0	8.4	4.8	9.4	5.4
1995	16.0	8.804	16.4	14.2	8.6	1.8	11.6	5.4
1996	20.4	11.852	18.4	13.6	8.8	1.8	10.6	8.8
1997	19.6	10.044	17.6	12.8	10.8	5.0	10.4	8.4
1998	20.4	14.004	18.2	14.8	10.4	7.4	8.2	9.4
1999	20.2	13.544	17.8	16.4	8.6	5.2	8.0	10.0
2000	16.2	14.656	14.4	15.6	9.6	8.4	8.6	9.8
2001	15.6	13.364	15.0	16.0	10.4	7.2	13.0	8.8
2002	19.2	14.024	16.6	15.6	11.2	9.0	12.0	7.4
2003	15.0	12.876	20.0	15.0	11.6	6.6	10.4	7.6
2004	16.0	11.344	19.6	15.0	10.2	7.6	10.0	8.0
2005	18.8	9.868	17.2	11.2	9.6	6.8	10.0	5.2
Mean	17.15	12.085	16.975	13.713	8.963	4.663	10	6.75
Rank	1	4	2	3	6	8	5	7

4.4.3 Correlations between S_h and SS_h and a Firm's Profits

Table 5 shows the rank correlations between yearly $S_h(Y)$ and yearly $SS_h(Y)$ and firms' profits for the eight pharmaceutical companies under study. The S_h and SS_h values refer to the years 1990–2005, and firms' profits refer to the period, 2006–2010. The Spearman rank correlation coefficient between the yearly S_h and a firm's profits is 0.857 (p = 0.007) when using $IPCh_4$, and is 0.762 (p = 0.028) calculated by $IPCh_3$. These results mean that the correlations can be described as "very strong". We note that $S_h(Y)$ based on $IPCh_4$ has the higher correlation with profits. Moreover, the Spearman rank correlation coefficient between the yearly SS_h (using $IPCh_4$) and a firm's profits is 0.810 (p = 0.015). This value can also be described as "very strong".

4.4.4 Differences of Yearly S_h and SS_h Indices of Firms between Two Different Profit Groups

Tables 6 and 7 present the results of a longitudinal observation combined with a nested case-control design. Obviously, the yearly S_h and SS_h indices of firms in

Table 5. Correlations among yearly S_h and yearly SS_h on the one hand and a firm's profits on the other.

Company	Profits 2006–2010 millions of US dollars (Average)	Rank profits 2006–2010	Yearly S_h Using $IPCh_4$		Yearly S_h Using $IPCh_3$		Yearly SS_h	
			Yearly S_h (Mean)	Rank S_h	Yearly S_h (Mean)	Rank S_h	Yearly SS_h (Mean)	Rank SS_h
Johnson & Johnson	11,451.00	1	234.31	2	170.19	2	17.15	1
Pfizer	10,461.00	2	162.63	3	101.25	4	12.09	4
Merck	6,610.02	3	246.19	1	175.19	1	16.98	2
Bristol-Myers Squibb	4,521.80	4	158.81	4	106.25	3	13.71	3
Amgen	3,718.20	5	90.06	6	53.94	6	8.96	6
Biogen Idec	554.10	6	35.69	8	24.36	8	4.66	8
Allergan	395.24	7	88.19	5	61.31	5	10.00	5
Genzyme	349.72	8	53.13	7	35.75	7	6.75	7
Spearman correlation				0.857**		0.762*		0.810*

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed).

Group HP are much bigger than those in Group LP during the period 1990–2005; these differences are significant. We note that the same statistical significances of paired differences are valid for results of S_h indices as well as for SS_h indices.

Table 6. Results of paired differences tests of firms' yearly S_h between Group H and Group L (based on $IPCh_4$).

Pairs- S_h	Paired differences Mean	Paired differences Std. Deviation	95% confidence interval of the difference Lower	95% confidence interval of the difference Upper	t-value	df	Sig. (2-tailed)
GHR1 - GLR1	144.250	45.918	119.782	168.718	12.566	15	0.000
GHR2 - GLR2	126.938	34.555	108.524	145.351	14.694	15	0.000
GHR3 - GLR3	158.000	46.286	133.336	182.664	13.654	15	0.000
GHR4 - GLP4	105.688	26.630	91.497	119.878	15.875	15	0.000

Table 7. Results of paired differences tests of firms' yearly SS_h between Group H and Group L (based on $IPCh_4$).

Pairs- SS_h	Paired differences Mean	Paired differences Std. Deviation	95% confidence interval of the difference Lower	95% confidence interval of the difference Upper	t-value	df	Sig. (2-tailed)
GHR1 - GLR1	8.188	2.331	6.94560	9.429	14.052	15	0.000
GHR2 - GLR2	7.422	2.700	5.98329	8.861	10.996	15	0.000
GHR3 - GLR3	6.975	2.213	5.79602	8.154	12.610	15	0.000
GHR4 - GLR4	6.963	1.570	6.12586	7.799	17.738	15	0.000

Figure 1 shows average profit values as a function of average $S_h(Y)$ values (using $IPCh_4$). As the Pearson correlation R is about 0.83, the rank correlation of Table 2 as well as the results shown in Table 5 are logical consequences of this relation. Note that, although this figure consists of just eight points, each of them is the result of thousands of values.

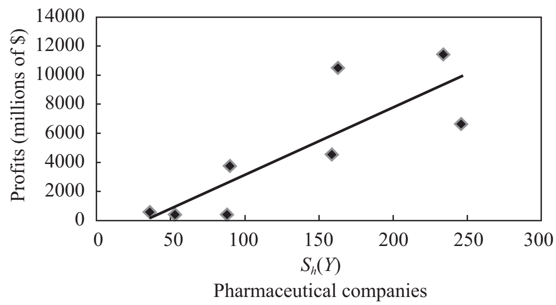

Figure 1. Functional relation between the $S_h(Y)$ values and profits.

5 Discussion and Conclusions

In many scientific fields, it is difficult to collect large samples to perform an "ideal" real-world investigation. Therefore, special approaches are developed and carefully designed for small samples. In this contribution we included a nested case-control approach, a method often used in the medical sciences, and applied it to improve the methodology used in patent research. By way of discussion we address the following issues.

5.1 The New $S_h(Y)$ Index Indicates the Primary Competitive Strength of a Firm's Patent Portfolio

Compared to the case of $IPCh_3$, S_h based on $IPCh_4$ can better indicate a firm's innovative activities, measured through patents, as well as their technological breadth, and map the potential market value of patents. Instead of the yearly h-indices which may represent a firm's innovation activities and their influence, the $S_h(Y)$ index, proposed in our investigation, can reflect a firm's innovation activities, its technological breadth, and its influence in an integrated way. As such the new index reflects the primary structure of a firm's patents and their influence and is an indicator for the "weight" related to primary competitive strength of a firm's patent portfolio (with significant correlation to a firm's profits).

5.2 The Breadth of Patent is a Primary Structure Affecting Its Performance

Although $SS_h(Y)$ is a comprehensive indicator for the "weight" of the essential, competitive strength of a firm's patent portfolio (including the depth of patents), and although the relation between $SS_h(Y)$ and a firm's profits is also significant, it does not have the same "strong" correlation as the $S_h(Y)$ index does, which suggests that the breadth of a patent is the primary structure affecting a patent performance. The depth of a patent plays a smaller role in a firm's profit. The Spearman rank correlation coefficient between the yearly Dh and a firm's profits is 0.690 (p = 0.058), while this correlation between the yearly average depth of patents and a firm's profits is -0.024, and hence is not significant (Appendix Tables A5 and A7).

5.3 The h-core Reflects Market Value and Technological Value

The first h items in a firm's patent list, known as its h-core, reflect market value and technological value. These core patents are closely related to the competitive strength of a company. Although there are multiple dimensions involved in the innovative performance of a firm, the core competitive strength of a company is highly dependent on the performance of patents (Hagedoorn & Cloodt, 2003), one aspect being that patents are transferable, so that the patent assignee benefits in monetary terms from their purchase (Lee, 2009; Palokangas, 2011).

Our work further leads to the suggestion to different sized firms to include policy-making on technological innovation in its management. This is because there is always a limited R&D capital in a company. Indeed, we also found out that the Spearman correlation coefficient between the yearly average number of 4-digit codes of patents and a firm's profits is even negative (namely –0.310, Appendix Tables A6 and A7), suggesting that a firm's profits are highly dependent on the first h items of a firm's patents rather than the "average patent" (Palokangas, 2011; Reitzig, 2003). The fact that a small group of patents essentially determines the competitive strength of a company is yet another example of the law of the vital few, also known as the 80–20 rule. In this sense, we claim that the structural h-index proposed in this study will be beneficial for modelling an optimal patent system.

Patent evaluation is a complicated issue which requires taking a full picture from different perspectives. This preliminary study proposes a new and simple indicator for gauging a company's patent portfolio. Positive results are backed by evidence based on a large dataset from the pharmaceutical industry. Of course, we are aware that this is just a case study and, moreover, that any R&D indicator is at best PAC, as put forward in the case of citation indicators by Rousseau (2016). We are convinced though that the structural h-index is a useful addition to the field of patentometrics.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant Nos: 71173185 and 71573225). The authors thank Jiancheng Guan (Professor in innovation policy at the University of the Chinese Academy of Sciences) and Yi Shen (Professor in medical statistics at Zhejiang University) for helpful discussions. We also thank Chaoyang Zhu for programming and Hui Jiang for help in data collection. Finally, we thank the reviewers for helpful suggestions leading to an improved manuscript.

Author Contributions

X. J. Hu (xjhu@zju.edu.cn) proposed the research idea, planned and designed the outline, carried out the data collection and data analysis, and wrote the first draft. R. Rousseau (ronald.rousseau@kuleuven.be, corresponding author) revised the plan and outline, joined discussion of the findings and contributed to writing the paper and its revision after review.

References

Banerjee, P.M., & Cole, B. M. (2010). Breadth-of-impact frontier: How firm-level decisions and selection environment dynamics generate boundary-spanning inventions. Technovation, 30(7), 411–419.

Banks, M.G. (2006). An extension of the Hirsch index: Indexing scientific topics and compounds. Scientometrics, 69(1), 161–168.

Bornehag, C.G., Sundell, J., Weschler, C.J., Sigsgaard, T., Lundgren, B., Hasselgren, M., & Hägerhed-Engman, L. (2004). The association between asthma and allergic symptoms in children and phthalates in house dust: A nested case-control study. Environmental Health Perspectives, 112(14), 1393–1397.

Bouyssou, D., & Marchant, T. (2011). Ranking scientists and departments in a consistent manner. Journal of the American Society for Information Science and Technology, 62(9), 1761–1769.

Braun, T., Glänzel, W., & Schubert, A. (2005). A Hirsch-type index for journals. The Scientist, 19(22), 8.

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

Chen, J.H., Jang, S.L., & Wen, S.H. (2010). Measuring technological diversification: Identifying the effects of patent scale and patent scope. Scientometrics, 84(1), 265–275.

Chen, Y.S., Shih, C.Y., & Chang, C.H. (2013). Patents and market value in the U.S. pharmaceutical industry: new evidence from threshold regression. Scientometrics, 97(2), 161–176.

Chiu, Y.C., Lai, H.C., Liaw, Y.C., & Lee, T.Y. (2010). Technological scope: Diversified or specialized. Scientometrics, 82(1), 37–58.

Denicolò, V. (1996). Patent races and optimal patent breadth and length. Journal of Industrial Economics, 44(3), 249–265.

Gilbert, R., & Shapiro, C. (1990). Optimal patent length and breadth. RAND Journal of Economics, 21(1), 106–112.

Goldstein, L., & Zhang, H.M. (2009). Efficiency of the maximum partial likelihood estimator for nested case control sampling. Bernoulli, 15(2), 569–597.

Grönqvist, C. (2009). The private value of patents by patent characteristics: Evidence from Finland. Journal of Technology Transfer, 34(2), 159–168.

Guan, J.C., & Gao, X. (2009). Exploring the h-index at patent level. Journal of the American Society for Information Science and Technology, 60(1), 35–40.

Hagedoorn, J., & Cloodt, M. (2003). Measuring innovative performance: Is there an advantage in using multiple indicators? Research Policy, 32(8), 1365–1379.

Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy Sciences of the United States of America, 102(46), 16569–16572.

Hu, X.J., & Rousseau, R. (2015). A simple approach to describe a company's innovative activities and their technological breadth. Scientometrics, 102(1), 1401–1411.

Hu, X.J., Rousseau, R., & Chen, J. (2012). A new approach for measuring the value of patents based on structural indicators for ego patent citation networks. Journal of the American Society for Information Science and Technology, 63(9), 1834–1842.

ISTIS: Institute of Scientific & Technical Information of Shanghai (2003). Feature analysis on global pharmaceutical industry 2002–2003, small change in the periodic R&D. Retrieved on August 20, 2012, from http://www.istis.sh.cn/list/list.aspx?id=3958 (in Chinese).

Kanniainen, V., & Stenbacka, R. (2000). Endogenous imitation and implications for technology policy. Journal of Institutional and Theoretical Economics, 156(2), 360–381.

Klemperer, P. (1990). How broad should the scope of patent protection be? RAND Journal of Economics, 21(1), 113–130.

Lee, Y.G. (2009). What affects a patent's value? An analysis of variables that affect technological, direct economic, and indirect economic value: An exploratory conceptual approach. Scientometrics, 79(3), 627–637.

Lerner, J. (1994). The importance of patent scope: an empirical analysis. RAND Journal of Economics, 25(2), 319–333.

Leydesdorff, L. (2015). Can technology life-cycles be indicated by diversity in patent classifications? The crucial role of variety. Scientometrics, 105(3), 1441–1451.

Liu, Y.X., & Rousseau, R. (2009). Properties of Hirsch-type indices: The case of library classification categories. Scientometrics, 79(2), 235–248.

Lodh, S., & Battaggion, M.R. (2014). Technological breadth and depth of knowledge in innovation: The role of mergers and acquisitions in biotech. Industrial and Corporate Change, 24(2), 383–415.

McMillan, G.S., & Thomas, P. (2005). Financial success in biotechnology: Company age versus company science. Technovation, 25(5), 463–468.

Merges, R.P., & Nelson, R.R. (1990). On the complex economics of patent scope. Columbia Law Review, 90(4), 839–916.

Miller, D.J. (2006). Technological diversity, related diversification, and firm performance. Strategic Management Journal, 27(7), 601–619.

O'Donoghue, T., Scotchmer, S., & Thisse, J.F. (1998). Patent breadth, patent life, and the pace of technological progress. Journal of Economics & Management Strategy, 7(1), 1–32.

Olivo, C., Lebedeva, I., Chu, C.Y., Lin, C.Y., & Wu, S.Y. (2011). A patent analysis on advanced biohydrogen technology development and commercialisation: Scope and competitiveness. International Journal of Hydrogen Energy, 36(21), 14103–14110.

Ozman, M. (2007). Breadth and depth of main technology fields: An empirical investigation using patent data. Middle East Technical University Working Paper, 2007.

Palokangas, T. (2011). Optimal patent length and breadth in an economy with creative destruction and non-diversifiable risk. Journal of Economics, 102(1), 1–27.

Prencipe, A. (2000). Breadth and depth of technological capabilities in CoPS: The case of the aircraft engine control system. Research Policy, 29(7–8), 895–911.

Reitzig, M. (2003). What determines patent value? Insights from the semiconductor industry. Research Policy, 32(1), 13–16.

Rousseau, R. (2016). Citation data as a proxy for quality or scientific influence are at best PAC (Probably Approximately Correct). Journal of the Association for Information Science and Technology (to appear); DOI: 10.1002/asi.23525.

Sapsalis, E, van Pottelsberghe de la Potterie, B., & Navon, R. (2006). Academic versus industry patenting: An in-depth analysis of what determines patent value. Research Policy, 35(10), 1631–1645.

Trajtenberg, M. (1990). A penny for your quotes: Patent citations and the value of inventions. Rand Journal of Economics, 21(1), 172–187.

Valiant, L. (2013). Probably Approximately Correct. New York: Basic Books.

Waltman, L., & van Eck, N.J. (2012). The inconsistency of the h-index. Journal of the American Society for Information Science and Technology, 63(2), 406–415.

Wang, Q., & von Tunzelmann, N. (2000). Complexity and the functions of the firm: Breadth and depth. Research Policy, 29(7–8), 805–818.

WIPO (2000). Patent law treaty, available at http://www.wipo.int/treaties/en/ip/plt/ (last visited on 2016, February 20).

Zhang, G.P., Chen, X.D., & Niu, X. (2012). The technology complexity based on patent width and depth (in Chinese). Science Research Management, 33(3), 113–135.

Appendix A

Table A1. Yearly h-indices for eight companies during the period 1990–2005.

Year	Johnson & Johnson	Pfizer	Merck	Bristol-Myers Squibb	Amgen	Biogen Idec	Allergan	Genzyme
1990	32	23	25	17	8	1	11	4
1991	29	24	31	23	12	0	20	4
1992	29	27	32	20	10	1	20	5
1993	32	22	32	25	7	1	19	4
1994	31	25	30	26	13	6	17	7
1995	30	18	30	27	14	2	23	7
1996	41	24	35	26	14	2	21	14
1997	39	20	33	23	16	6	19	14
1998	40	29	35	29	15	11	14	15
1999	40	28	34	32	13	7	14	16
2000	29	30	25	29	14	14	16	15
2001	27	26	26	31	12	10	26	13
2002	39	28	30	26	14	13	23	10
2003	26	25	37	27	17	8	18	10
2004	30	22	37	27	15	11	18	11
2005	36	18	30	19	13	8	17	5

Table A2. Yearly $IPCh_4$ for eight companies during the period 1990–2005.

Year	Johnson & Johnson	Pfizer	Merck	Bristol-Myers Squibb	Amgen	Biogen Idec	Allergan	Genzyme
1990	6	7	6	5	4	1	4	3
1991	6	6	6	6	6	0	5	3
1992	5	6	7	5	4	1	5	4
1993	7	6	8	5	5	1	4	3
1994	7	6	8	7	6	5	5	5
1995	8	6	8	6	6	2	4	5
1996	8	5	8	6	6	2	4	6
1997	8	7	8	6	8	5	5	5
1998	7	6	8	6	8	6	5	6
1999	6	7	8	6	6	5	4	6
2000	7	8	8	7	7	5	4	6
2001	7	8	8	6	10	6	5	6
2002	7	9	8	8	10	7	5	6
2003	9	7	9	7	8	7	6	6
2004	7	6	8	7	7	6	5	6
2005	8	6	9	6	7	7	6	6

Table A3. Yearly $IPCh_3$ for eight companies during the period 1990–2005.

Year	Johnson & Johnson	Pfizer	Merck	Bristol-Myers Squibb	Amgen	Biogen Idec	Allergan	Genzyme
1990	5	4	4	4	3	1	3	2
1991	5	5	5	4	3	0	4	3
1992	4	4	5	4	3	1	4	3
1993	5	4	6	4	3	1	3	2
1994	5	4	6	4	3	3	3	3
1995	5	3	5	4	4	2	3	3
1996	6	4	5	4	4	2	3	4
1997	6	4	6	4	4	4	3	4
1998	5	4	5	4	5	4	3	4
1999	4	4	6	4	4	3	3	4
2000	5	5	6	4	5	4	3	4
2001	5	4	5	4	5	4	3	4
2002	5	5	6	5	5	4	3	4
2003	6	4	6	5	5	5	4	4
2004	5	4	6	4	4	4	4	4
2005	6	4	7	5	5	5	4	3

Table A4. The yearly h-index of patent depth (Dh_y) for eight companies during the period 1990–2005.

Year	Johnson & Johnson	Pfizer	Merck	Bristol-Myers Squibb	Amgen	Biogen Idec	Allergan	Genzyme
1990	3	6	5	3	3	1	2	2
1991	4	6	5	4	3	0	4	2
1992	4	5	6	3	4	1	3	3
1993	3	5	6	4	4	1	4	2
1994	3	6	10	4	4	2	3	3
1995	4	4	6	5	3	1	4	3
1996	4	6	6	4	4	1	3	4
1997	4	5	6	6	6	3	4	4
1998	8	6	5	4	6	3	3	5
1999	9	6	5	6	5	2	4	6
2000	9	7	6	6	6	4	3	7
2001	10	8	7	6	8	4	3	6
2002	4	8	7	10	8	5	4	5
2003	5	8	8	7	8	3	4	6
2004	6	7	8	7	7	4	4	6
2005	6	8	8	6	8	4	4	4

Table A5. Yearly average depth of patents (average d_{ad}) for eight companies during the period 1990–2005.

Year	Johnson & Johnson	Pfizer	Merck	Bristol-Myers Squibb	Amgen	Biogen Idec	Allergan	Genzyme
1990	1.50	2.49	1.79	1.77	2.04	4.22	1.44	2.29
1991	1.72	2.53	1.67	1.76	2.81	0.00	1.80	2.13
1992	1.59	2.35	1.93	1.67	2.86	2.83	1.48	2.39
1993	1.63	2.59	1.86	1.79	3.94	2.28	1.79	2.52
1994	1.52	2.47	2.58	1.70	3.71	2.53	1.62	2.00
1995	1.80	2.01	2.01	2.04	2.10	1.66	1.74	2.01
1996	1.78	2.63	2.01	1.75	2.29	3.63	1.66	2.85
1997	1.67	2.61	1.95	1.82	2.65	2.89	2.07	2.40
1998	2.22	3.01	2.01	1.73	3.15	2.82	1.53	3.18
1999	2.24	2.86	2.08	2.12	2.89	2.41	2.17	2.72
2000	2.38	3.14	2.22	2.40	2.93	3.33	1.68	3.44
2001	2.74	3.41	2.26	2.56	2.47	3.54	1.68	3.05
2002	1.69	3.06	2.33	3.34	3.65	3.18	2.03	2.30
2003	1.77	3.19	2.46	2.73	3.50	2.42	1.81	2.69
2004	1.93	2.86	2.65	2.48	3.99	2.78	1.99	2.57
2005	2.016	2.67	2.64	2.31	3.96	3.01	2.00	2.01
Mean	1.887	2.743	2.153	2.123	3.059	2.721	1.781	2.543
Rank	7	2	5	6	1	3	8	4

Table A6. Yearly average number of 4-digit IPC codes (ave IPC-4 codes) of patents for eight companies during the period 1990–2005.

Year	Johnson & Johnson	Pfizer	Merck	Bristol-Myers Squibb	Amgen	Biogen Idec	Allergan	Genzyme
1990	2.85	2.81	2.83	2.76	3.46	9.00	2.58	2.67
1991	3.14	2.92	2.95	2.86	4.47	0.00	2.51	4.00
1992	2.37	2.60	3.00	2.72	3.63	6.00	2.84	3.91
1993	2.99	2.86	2.75	2.75	4.08	7.00	1.97	2.55
1994	2.98	2.75	2.72	2.97	4.17	6.00	2.54	4.07
1995	3.19	2.89	3.06	3.02	4.39	7.50	2.50	3.50
1996	3.52	2.94	3.07	2.83	4.08	6.50	1.94	3.57
1997	3.43	3.50	2.94	2.74	4.92	6.60	2.50	3.67
1998	2.60	3.60	3.12	2.59	4.64	5.56	2.36	3.08
1999	2.33	3.62	3.31	2.61	3.88	4.36	2.43	3.23
2000	2.96	3.78	3.47	3.14	4.08	4.57	2.18	3.49
2001	3.07	3.78	3.52	3.08	5.62	4.72	2.16	3.17
2002	2.78	3.72	3.74	3.55	6.06	4.74	2.43	3.57
2003	2.26	3.42	3.67	3.33	4.99	5.91	2.71	3.12
2004	2.91	2.97	3.48	3.04	4.66	4.51	2.56	3.29
2005	2.91	2.70	3.76	3.02	3.93	4.82	2.76	3.42
Mean	2.89	3.18	3.21	2.94	4.44	5.49	2.44	3.39
Rank	7	5	4	6	2	1	8	3

Table A7. Correlations among yearly average IPC-4 codes and yearly average d_{ad} of patents and a firm's profits.

Company	Rank profits 2006–2010	Yearly average IPC-4 codes		Yearly average d_{ad}	
		Mean	Rank	Mean	Rank
Johnson & Johnson	1	2.89	7	1.887	7
Pfizer	2	3.18	5	2.743	2
Merck	3	3.21	4	2.153	5
Bristol-Myers Squibb	4	2.94	6	2.123	6
Amgen	5	4.44	2	3.059	1
Biogen Idec	6	5.49	1	2.721	3
Allergan	7	2.44	8	1.781	8
Genzyme	8	3.39	3	2.543	4
Spearman correlation			–0.310		–0.024