Patent Citation Analysis: Calculating Science Linkage Based on Citing Motivation

Rui Li
Department of Information Management, School of Public Administration, Sichuan University, 29 Wangjiang Road, Chengdu, Sichuan 610064, P.R. China. E-mail: li_rui@scu.edu.cn

Tamy Chambers, Ying Ding, and Guo Zhang
Department of Information and Library Science, School of Informatics and Computing, Indiana University, 1320 East 10th Street, LI 011, Bloomington, Indiana 47405-3907. E-mail: {tischt,dingying,guozhang}@indiana.edu

Liansheng Meng
National Science Library, Chinese Academy of Sciences, 33 Beisihuan Xilu, Zhongguancun, Beijing 100190, P.R. China. E-mail: menglsh@mail.las.ac.cn

Science linkage **is a widely used patent bibliometric indicator to measure patent linkage to scientific research based on the frequency of citations to scientific papers within the patent. Science linkage is also regarded as noisy because the subject of patent citation behavior varies from inventors/applicants to examiners. In order to identify and ultimately reduce this noise, we analyzed the different citing motivations of examiners and inventors/applicants. We built 4 hypotheses based upon our study of patent law, the unique economic nature of a patent, and a patent citation's market effect. To test our hypotheses, we conducted an expert survey based on our science linkage calculation in the domain of catalyst from U.S. patent data (2006–2009) over 3 types of citations: self-citation by inventor/applicant, non-self-citation by inventor/applicant, and citation by examiner. According to our results, evaluated by domain experts, we conclude that the non-self-citation by inventor/applicant is quite noisy and cannot indicate science linkage and that self-citation by inventor/applicant, although limited, is more appropriate for understanding science linkage.**

Introduction

The most valuable scientific researchers are those who produce critical technologies key to the further advancement and success of human society. In line with current initiatives in translational science (Owen-Smith & Powell, 2003), there is pressure to shorten the distance between scientific research and societal application. Establishing and understanding the complex relationship between science research and technological invention continues to be of interest to scholars and scientists, as well as to governments, institutions, and funding agencies. *Science linkage* is a helpful indicator for discovering the value of scientific research and its place in forecasting future critical and key technology. It calculates the coupling and collaboration effect between scientific research and technological invention. For scholars and scientists, science linkage is meaningful because it provides empirical evidence to illustrate the relationship between science and technology. For governments, institutions, and funding agencies, science linkage is meaningful as it provides useful reference for science and technology policy, science–technology integration, and funding decisions.

Patents can play a key role in understanding the link between scientific research and technology innovation or discovery and application. However, as others have pointed out (e.g., Jaffe, Fogarty, & Banks, 1998; Jaffe, Trajtenberg, & Fogarty, 2000), understanding this link can be difficult because patents serve multiple functions and can be defined differently from various perspectives. In turn, if we use citation analysis, each perspective will also yield different citation motivations, thus making the analysis even more complicated. For example, if we view a patent as a type of specification document, then papers cited in the

Received April 9, 2013; revised June 17, 2013; accepted June 18, 2013

• Published online 22 January 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.23054

patent could be analyzed similar to those cited in journal articles. However, if we view the patent as a legal document, which defines rights and focuses on the patent's claims, then papers cited in the patent would carry specific legal functions prescribed by patent law. We could also view the patent as a type of economic interest document, which describes the product's benefits versus competitors and its potential marketability. In that case, the papers cited would most likely work conversely from those cited in journals—criticizing instead of giving credit (Weinstock, 1970).

Citations within patents are often referred to as "prior art" and usually reference either other patents or non-patent references, mostly scientific literature. Although extensive work has been done on patent reference citation as a means of establishing science linkage (e.g., Criscuolo & Verspagen, 2008; Jaffe et al., 2000), in this article we will use non-patent references, pioneered by Narin and Noma (1985), to evaluate science linkage. According to Harhoff, Scherer, and Vopel (2003), scientific papers account for about 60% of all non-patent references and these references can be added to a patent application by either the inventor/applicant or the government examiner of that patent. As of 2006, all references in USPTO (United States Patent and Trademark Office) patents have been marked as either *cited by examiner* or *cited by others*, which now allows us to evaluate the science linkage of each type of citation: inventor/applicant and examiner. Since patents also carry an economic component, which leads to heavy criticism of competitive technologies within the patents, inventor/applicant citations have often been identified as noisy with regard to science linkage (Hall, Jaffe, & Trajtenberg, 2005; Jaffe et al., 2000). For this reason, we will further divide the inventor/applicant group into self-cited and non-self-cited citations to achieve a better analysis.

It is our contention that whether or not scientific papers cited in patents indicate science linkage depends on who cites them (inventor/applicant or examiner), why they are cited (inventor/applicant application or examination process), and how they are cited (self-citation by inventor/applicant or non-self-citation by inventor/applicant). In this article, we will examine the link between science and technology by analyzing how the citing motivation of both inventor/applicant and examiner impacts science linkage. To do so, we will first identify the difference between inventor/applicant self-citation and non-self-citation motivation compared with examiner citation motivation. We will then calculate the science linkage of each set and ultimately propose a better way of identifying and evaluating the science linkage between science and technology using non-patent citation.

The remainder of this article is organized as follows. In the first section, we review the literature and examine the current conflictive use of scientific literature in patent analysis. In the section titled Motivation Difference Between Patent Citation and Journal Article Citation, we analyze the difference between the purpose and use of scientific paper citation in academic literature and patents; and in the following section (Motivation Difference Between Examiner and Inventor/Applicant) we describe the citation motivations of the three types of citations: inventor/applicant self-cited, examiner, inventor/applicant non-self-cited. In the next three sections, we propose our hypotheses regarding citation motivations in relation to science linkage, describe our proposed test methods, and outline and identify the development of our data set. Then, in the Experimental Calculation of Science Linkage section, we describe our process of calculating science linkage. The Judgment by Domain Experts section describes the expert survey and provides analysis of their responses. In the final three sections, we examine and test of our hypotheses based on the results, discuss the results, and provide concluding remarks and recommendations.

Literature Review

Science linkage, as defined by Narin and his colleagues (Narin, Hamilton, & Olivastro, 1997; Narin, Hamilton, & Olivastro, 1995) at Computer Horizons Inc. (CHI), has been widely used as a patent bibliometric indicator to measure patent linkage/citation to scientific research. Science linkage is usually quantified as the total scientific papers cited in a patent. A growing number of researchers have applied patent citation techniques to analyze scientific innovations (Bacchiocchi & Montobbio, 2009; Bhattacharya, Kretschmer, & Meyer, 2003; Chen & Hicks, 2004; Hu, Chen, Huang, & Roco, 2007; Verbeek, Debackere, & Luwel, 2003; Wong & Ho, 2007). However, there has been much debate over the last few years regarding the interpretation of scientific papers cited in patents. Many researchers (Carpenter & Narin, 1983; Grupp, 1996; Narin et al., 1995; Rip, 1992) believe the embedded knowledge of scientific papers cited in patents indicates prior usage in the development of these patents. Therefore, they contend these papers serve the same function as papers cited in journal articles and would thus be considered a linkage between technology and science. However, others (Breschi & Catalini, 2010; Chen & Hicks, 2004; Jaffe et al., 2000; Tijssen, 2001) have argued that patent citations can be interpreted in various ways, including some which do not point to the actual flow of knowledge from cited to citing. These researchers also introduce the concept of a circular versus linear flow between science research and technological innovation, which again is consistent with current translational science studies. Similarly, Meyer (2000) ultimately concluded it was risky to count scientific research papers cited in patents as contributions from science to technology and Karvonen and Kässi (2013) concluded that non-patent literature produced ambiguous results with unclear validity.

In fact, patent citation behavior is extremely complex because of multiple citers within the same patent: the

DOI: 10.1002/asi

patent examiner and the inventor/applicant. Azagra-Caro, Mattsson, and Perruchas (2011) contend that examiner citations are for the purpose of restricting patent claims, whereas inventor/applicant citations are for demonstrating prior work/art related to the invention. Lai and Wu (2005) claim that patent examiners, as government agents who approve patent applications, produce more credible citations and that more effort should therefore be devoted to better codification of patent citations. Stock and Stock (2006) further propose that examiner citations should be used to build indicators such as h-indexes of firms. However, Meyer (2000) argues that examiner citations may provide biased information about knowledge flows based on non-technoeconomic reasons, such as examiner workload, claim volume, a duty of disclosure, patent examiner education, office methods, and a preference for national or English language.

Conversely, others (Kesan, 2002; Thompson, 2006; Alcácer, Gittelman, & Sampat, 2009) have noted that since inventors are more familiar with their inventions than examiners, their role in determining science linkage is crucial. Criscuolo and Verspagen (2008) argue that the knowledge base of a patent will appear to be more localized if measured through inventor/applicant citations. Similarly, Azagra-Caro, Fernández-de-Lucio, Perruchas, and Mattsson (2009) contend the degree of localization and differences between examiner and inventor/applicant citations depend on the absorptive capacity, and thus highlight the use of inventor/applicant rather than examiner citations as a better expression of knowledge flows.

Alcácer and Gittelman's case study (2006), however, violated this assumption by showing that examiner citations were more localized than inventor/applicant citations in real-world cases, contrary to the expectation that inventor/applicants preferentially cite proximate technologies. Hall et al. (2005) explained the competitive use of citations as follows: If patent B cites the prior work/art of A, it implies that A represents a piece of previously existing knowledge upon which B is built and over which B cannot have a claim. Hence, Lampe (2012) explicitly analyzed an applicant's citation as a strategic decision. He concluded that if an inventor/applicant omits a citation to A, then he/she can potentially claim ownership over that technology embodied in A, and that this ownership claim may entitle the inventor/applicant to royalty payments from competing firms. Using a sample of 267 patent lawsuits, Allison and Lemley (1998) found that the probability of invalidity, based on cited prior work/art, was 30% compared to 41% for uncited prior work/art, indicating that some closely related prior work/art may have been withheld by inventor/applicant.

In summary, previous research has introduced two conflicting implications. First is the contention that scientific papers cited by examiners are more creditable because examiners, as government agents responding to patent law, cite prior work/art comprehensively, whereas inventor/applicants omit/withhold prior work/art strategically to gain an economic interest. The second contention is that scientific papers cited by an inventor/applicant are more reliable because, as an inventor is more familiar with his/her invention, citations will be more localized, whereas examiners may provide biased information.

Because of these two competing contentions, the effect of examiner or inventor/applicant citations on a patent's science linkage calculation is ambiguous. Additionally, few studies have investigated the diversity of an inventor/applicant's citing behavior and we have no insight regarding an inventor/applicant's complex citing motivations.

Motivation Difference Between Patent Citation and Journal Article Citation

The inventor/applicant citations to scientific papers within a patent are different from those by an author within a journal article because of the ultimate goal of the citation, the social effects of the citation, and the legal functions of the citation. The following will review each in detail.

Ultimate Goals

Because the ultimate goal of science is to seek truth, journal article authors, as scientists, tend to offer a realistic description about the surrounding world. They seek to answer the questions "What is it?" and/or "Why is it?" Their ultimate goal, by citing scientific papers, is to draw upon pioneer scientific research achievements and to share useful knowledge with newcomers. Different from science, however, the ultimate goal for inventors/applicants in a patent application is to assert uniqueness/betterness. Patent inventors/applicants offer new products to improve life quality by seeking to answer the questions "What to do instead?" and/or "How to do it better?" As patent law stipulates that any new patent should not directly use any prior work/art (mainly including prior patents and scientific papers, which are named as public knowledge), the inventor/applicant's ultimate goal for citing scientific articles becomes avoiding linkage for his or her products to public domain knowledge. Their goals are therefore opposite; patents (and by association their citations) seek market share through division with others, while journal articles (and their included citations) seek scholarly communication though unification with others.

Social Effects

Because they seek the truth, journal citations usually function as indicators of the consistent progress of science accumulation by recording knowledge diffusion across different domains. However, as a result of seeking uniqueness/improvement, patent citations represent what Schumpeter (1942) calls "creative destruction." This means that

DOI: 10.1002/asi

enterprises win profits through introducing new products and technologies to replace old ones. Caballero and Jaffe (1993) define patent citation as a parameter for the obsolescence of technology, with the *obsolescence rate* described by the patent citation function:

$$\alpha^*(t,s) \equiv \frac{C_{t,s}}{S_t P_s}$$

where $C_{t,s}$ refers to the frequency of patents at time point t which cite prior work/art at time point s. S_t is the number of sample patents at time point t, P_s is the number of prior work/art at time point s, and $\alpha^*(t, s)$ represents the depreciation of the prior work/art at time point s because of the new patents at time point t. In their work they estimated the depreciation of a set of observations consisting of (s, t) pairs with t varying between 1975 and 1992 and s varying between 1900 and t. Based on a data set of U.S. patents between 1975 and 1992, they calculated an estimated obsolescence rate of about 0.075 per year. Therefore, the social effect of citing scientific papers within articles is to share wisdom in a win-win game, whereas inventor/applicant scientific paper citing within patents is a fight for profits in a zero-sum game.

Legal Function

In a patent, references are required by patent law, while author citations within journal articles are not legally bounded behavior. According to the article 1104 of USPL (United States Patent Law) and chapter 6 of EPC (European Patent Convention), a patent inventor/applicant must cite prior work/art of the same subject to demonstrate the advancement of his/her new technology. Article 2257 of USPL additionally regulates the citing format. According to chapter 3 of JPL (Japanese Patent Law) and article 18 of CPLIR (China's Patent Law Implementing Rules), an inventor/applicant is required to cite prior work/art to describe the theoretical framework or technical background of his/her invention. Collins and Wyatt (1988) have summarized the legal function of inventor/applicant citing papers as "the inventor must set out the background in such a way as to show how the claimed invention relates to, but is innovatively different from what was already public knowledge, and his/her task is to identify his/her work either related to but significantly different from, or else a useful step towards an new invention or an use of the invention" (p. 66).

Motivation Difference Between Examiner and Inventor/Applicant

Inventor/Applicant Non-Self-Citing Motivation

It is widely believed that technical invention is related to, or in some cases initiated and/or stimulated by, scientific research activities performed in a related field. Therefore, the level of scientific paper citing is an appropriate proxy for quantifying the linkage between the technology field and the science domain (Schmoch, 1993). Due to the controlled nature of the patenting process and its legal consequence, non-self-citations by the inventor/applicant result from the required "search for prior-art." Patent inventors/applicants are subject to "duty of disclosure," which obliges them to disclose any relevant documents that might have a bearing on the patent claims. This legal requirement, combined with USPTO's rigorous enforcement of the disclosure of prior work/art, has motivated inventors/applicants to limit non-self-citing behavior to *description theory background* and *explanation of knowledge source*. According to Jaffe, Trajtenberg, and Henderson (1993), prior art cited in a patent application might cause rejection because it works as comparative public knowledge which might overthrow the novelty of the patent application. Hall (2000) and Criscuolo and Verspagen (2008) claimed that inventors/applicants strategically cite prior work/art, and Alcácer et al. (2009) argued that they might omit relevant information on purpose to avoid competitors.

Besides novelty, creativity is another legal requirement of a patent application. This means the inventor/applicant must identify his or her work as related to, but significantly different from, prior work/art or identify a creative use for the prior invention. With this in mind, an inventor/applicant might tend to cite disadvantages or defects of prior works. For example, in U.S. Patent 7374930, a patent for gene technology treating diabetes mellitus, all scientific papers cited by the inventor/applicant were about insulin research. The inventor/applicant did not cite these papers to show the linkage between insulin research and gene technology, but rather to point out the defects and side effects of current insulin technology in treating diabetes mellitus to emphasize the advantages of his own invention.

In general, an inventor/applicant can be very strategic in deciding what and how much prior work/art to cite, since these citations may affect the novelty and creativity of the patent and thereby the rights granted by the patent. Hence, inventor/applicant non-self-citing motivation likely includes: (a) description of theory background and knowledge source; (b) attribution to highlight defects or disadvantages of prior work; and (c) concealment of public knowledge, including prior patents and scientific papers by others closely linked to the patent application.

Inventor/Applicant Self-Citing Motivation

An inventor/applicant who self-cites reflects his or her dual role as a scientific researcher and technical innovator. The survey by Tijssen (2002) on inventors/applicants showed that 79% of them cited their own scientific research achievements in their patents. Others (Balconi, Breschi, & Lissoni, 2004; Breschi, Lissoni, & Montobbio, 2007) analyzed Italian patents and found that self-citation exhibits significant linkage between technology innovation and basic research in Italy. In a study of more than 400 patents

DOI: 10.1002/asi

in the field of biomedicine in Belgium, Sapsalis, Van Pottelsberghe de la Potterie, and Navon (2006) found that patents with a high proportion of self-citation were often imbued with high technical value. Breschi and Catalini (2010) concluded that inventor/applicants who self-cite in their patent applications act as gatekeepers who span the gap between the scientific and technical research communities.

In general, as a scientific researcher, the inventor/applicant contributes social value to his or her intellectual achievement by publishing papers. Similarly, as a technical innovator, the inventor/applicant contributes economic value to his or her intellectual achievement by transforming scientific findings into technical patents. Therefore, inventor/applicant self-citing behavior serves to bridge scientific research and technical innovation, as well as transform social value to economic value.

Examiner Motivation

An examiner's motivation for citing scientific papers includes two facets: providing comparative literature for examining patentability and providing evidence for limiting the scope of the claim. Since a patent is a form of legal right, patent law and the related regulations of the patent examination system contribute to the examiner's motivation. Within the patent examination system, launched in the United States in 1790, a patent examiner is responsible for examining the patentability of an application, limiting the scope of patent rights, and publicizing the technical content. As discussed earlier, the ultimate goal of a patent is to seek market share; as a result, the major task of an examiner is to guarantee the novelty of patent and to avoid reusing public knowledge.

According to article 301of USPL, articles 90-92 of EPC, article 63 of JPL, and article 38 of CPLIR, citation/reference is used as the relevant information for patentability examination. Although, patentability requires novelty, creativity, improvability, practicality, and feasibility, novelty is clearly the core component. When considering the requirement of novelty, Sternitzke (2009) suggested an examiner looks for earlier literature which has the same (or almost the same) features as the patent application. Only if no relevant literature can question the novelty of the invention can the patent application be accepted.

Similarly, article 700 in MPEP (*Manual of Patent Examining Procedure*) requires patent examiners to list all cited literature on the first page of the specification. MPEP 2200, MPEP 2253, MPEP 2275, and MPEP 2287 dictate that during the examination process, if an examiner cannot confirm the role of the cited prior art in an application, the examiner is required to request further explanation from the applicant about the difference between the application and the cited literature. In addition, article 132 of USPL states an examiner has to explain reasons and provide evidence through citation of prior works to reject or reexamine an application.

The scope of a patent claim is limited by the concept scope and exact words in the claim. It is further defined by the market share scope. Claims of a patent include the independent claim, describing the maximum scope of the market share, and subordinate claim, limiting the partial features or changing the independent claim. According to article 112 of USPL, article 84 of EPC, articles 70-71 of JPL, and article 21 of CPLIR, if patent B cites literature A (e.g., prior patent, scientific paper) then the corresponding claim scope of patent B will be limited within the concept scope of literature A. Sternitzke (2009) studied 2,719 World Patents issued in 1996 and found that 45.3% of scientific papers cited by the examiner were used to judge novelty and creativity, while 30.2% were used to limit the scope of claims.

Hypotheses

Based on the preceding, we infer that scientific paper self-citation by inventor/applicant is the major clue to tracing knowledge diffusion from scientific research to technical innovation, and that scientific papers cited by an examiner are relevant literature which provides evidence for examining patentability and limiting claims. However, we believe a large portion of non-self-citations by an inventor/applicant tend to be irrelevant literature to widen the claim and/or increase the economic gain. To clarify how scientific papers cited in patents can be used to measure science linkage, we conducted an exploratory study to investigate the following hypotheses:

H1: Scientific papers cited in patents can be considered as a measure of science linkage.

H2a: Scientific papers self-cited by the inventor/applicant are the best measure of linkage between science and technology.

H2b: Scientific papers cited by the examiner rank second to indicate the linkage between science and technology.

H2c: Scientific papers non-self-cited by the inventor/applicant are noisy and weakly indicate the linkage between science and technology.

Test Method

To test the above hypotheses, we conducted an exploratory study using an appropriate data set (see next section). We calculated science linkage based on the identified citation patterns for inventor/applicant self-citation, examiner citation, and inventor/applicant non-self-citation by inviting domain experts to evaluate all three patterns based on their profound domain knowledge. We used the analysis of the expert survey to order the three patterns. This analysis and ordering provided an understanding of the science linkage between science and technology based patent citation of scientific papers and served as a test of our hypotheses.

DOI: 10.1002/asi

TABLE 1. Quantity of papers cited within patents by International Patent Classification.*

	A	B	C	D	E	F	G	H	TOTAL
Cited Papers	4,066	14,395	18,169	11,953	3,908	5,868	14,007	26,824	99,190
USPTO	3,205	10,110	12,834	7,248	3,409	3,450	8,998	18,217	67,471 (68%)
EPB	606	2,904	4,013	3,339	329	1,554	3,390	5,937	22,072 (22%)
WIPO	255	1,381	1,322	1,366	170	864	1,619	2,670	9,647 (10%)

Note. *A. Human necessities; B. Operation and transportation; C. Chemistry and metallurgy; D. Textile and paper manufacture; E. Fixed structure; F. Mechanical engineering, lighting, heating, weapon, and blasting; G. Physics; H. Electricity

TABLE 2. Distribution of cited papers within the IPC domains.

	Mean	Maximum	Standard deviation	Skewness
A	7.97	114	11.87	4.15
B	6.83	183	14.37	6.14
C	8.88	472	24.47	10.13
D	6.37	263	14.22	9.27
E	15.57	209	39.71	3.46
F	5.58	273	21.04	10.11
G	8.15	134	13.73	4.59
H	12.30	826	44.65	10.32

Note. A. Human necessities; B. Operation and transportation; C. Chemistry and metallurgy; D. Textile and paper manufacture; E. Fixed structure; F. Mechanical engineering, lighting, heating, weapon and blasting; G. Physics; H. Electricity

Data set

To determine the most appropriate data set we randomly selected 10,000 patents from each International Patent Classification (IPC) domain (A. Human necessities; B. Operation and transportation; C. Chemistry and metallurgy; D. Textile and paper manufacture; E. Fixed structure; F. Mechanical engineering, lighting, heating, weapon and blasting; G. Physics; H. Electricity) issued by the U.S. Patent and Trademark Office (USPTO), the European Patent Bureau (EPB), and World Intellectual Property Organization (WIPO) using the *Derwent Innovation Index* (DII). We then identified all patents that cited scientific papers and confirmed that 99,190 papers were cited by the 80,000 patents surveyed. Consistent with other studies (Callaert, Van Looy, Verbeek, Debackere, & Thijs, 2006) and in large part due to the USPTO's duty of candor which requires inventor/applicants provide all prior work/art documents which are in any way relevant to the invention, we found that 68% of these patents were issued by the USPTO, as shown in Table 1. Therefore, we chose to narrow our analysis to only USPTO patents. Statistical analysis, as detailed in Table 2, revealed that domains C, E, F, and H were not good sample sources because of their high standard deviation values and that domain D was not good because it had a high skewness. Comparing domains A, B, and G we found B had the largest number of cited papers and therefore chose the domain "Operation and transportation" as our experiment source.

Within this domain, we narrowed the topic to "catalyst" as it plays an important role in modern industry and is a shared concern of the academic world. We then queried the USPTO Patent Database (http://patft.uspto.gov/) using the following query: TTL/catalyst and ISD/1/1/2006->1/1/2007. This resulted in the discovery of 452 patents, which cited 2,652 scientific papers in total. Using the citation indicators, *cited by examiner* or *cited by others*, we determined that 271 of these papers were cited by the examiner and 2,381 were cited by the inventor/applicant. Similar to Breschi and Catalini (2010), we determined which papers were self-cited through manual reference comparison of the patent and paper authors. This resulted in 176 of the papers found to be self-cited by the inventor/applicant. Our three citation groups were thus: (a) 176 scientific papers self-cited by inventor/applicant, (b) 271 scientific papers cited by examiner, and (c) 2,205 scientific papers non-self-cited by inventor/applicant.

Experimental Calculation of Science Linkage

Initial review of the data revealed a huge difference in quantity between the three citation types. The maximum number of non-self-citations by inventor/applicant was 8.13 times that of the citations by examiner and 12.52 times that of the self-citations by inventor/applicant. The fewest self-citations by an inventor/applicant was less than 1/10 that of non-self-citations by the inventor/applicant. We believe the differences between the citation types are the result of the diversity and complexity of citation motivations. As the non-self-citing motivation of the inventor/applicant is the most complex, including such things as *description of theory background*, *explanation of knowledge source*, *identification of problems in current practice*, and *illustration of prior art defects or disadvantages,* it is expected that these citations would show diversity in content together with larger quantity. Because examiner citing behavior, as regulated by patent law, is relatively simple and confined to *providing comparative literature for examining novelty and for limiting scope of claims*, papers cited by them should be inevitably small in number. The inventor/applicant self-citing motivation, as a *bridge between scientific research and technical innovation*, requires the dual role as a scientific researcher and technical innovator and, as a result, is the smallest number in citation quantity.

Each paper was assigned a domain based on the journal subject category from the *Journal Citation Reports* (Thomson Reuters, 2011). Then, based on the these three

DOI: 10.1002/asi

TABLE 3. Domains identified by citation type.

Pattern (a): Inventor/applicant self-cited		Pattern (b): Examiner cited		Pattern (c): Inventor/applicant non-self-cited			
Domains	SL	Domains	SL	Domains	SL	Domains	SL
Materials	0.3	Elec Chem	0.21	Phys Chem	0.9	Fluid	0.06
Elec Chem	0.09	Materials	0.18	Materials	0.84	Chroma	0.06
Phys Chem	0.06	Zeolites	0.12	Physics	0.63	Synthesized fibre	0.06
Microporous	0.06	Chroma	0.09	Elec Chem	0.51	Hydrocarbon	0.06
Platinum	0.06	Physics	0.06	Organ Chem	0.42	Microcapsules	0.06
Organ Chem	0.03	Spectrum	0.06	Zeolites	0.42	MacromolChem	0.06
Tetrahedron	0.03	Organometal	0.06	Ceramics	0.39	Biostruct	0.06
Chrome	0.03	Microcapsules	0.06	Organometal	0.36	Genes cells	0.06
Molecular catalysis	0.03	Petroleum	0.06	Molecular catalysis	0.36	Vegoil	0.03
Thermal Engineering	0.03	Molecular catalysis	0.06	Petroleum	0.30	Spectrum	0.03
Ceramics	0.02	Organ Chem	0.03	Microporous	0.24	Silicon	0.03
Heterogeneous	0.02	Inorgan Chem	0.03	Carbon	0.24	Tombar thite	0.03
Distil	0.01	Optics	0.03	Bioenergy	0.23	Nickel	0.03
		Polymer	0.03	Solid state ionics	0.21	Heteroatomic ring	0.03
		Mixedmetal Oxidation	0.03	Hydrogen energy	0.21	Mineral	0.03
		Microporous	0.03	Autoengineering	0.21	Environmental	0.03
		Nickel	0.03	Biotechniques	0.21	Pyrolysis	0.03
		Metallosilicates	0.03	Alloy	0.18	Oil&Gas	0.03
		Heteroatomic ring	0.03	Tetrahedron	0.18	Colloid	0.03
		Heteroatomic ring	0.03	Fuel battery	0.18	Separation	0.03
		Solid state ionics	0.03	Bio Chem	0.18	Sensor	0.03
		Fuel	0.03	Microwave	0.15	Protein	0.03
		Carbon	0.03	Polymer	0.15	Nucleic acids	0.03
		Clay	0.03	Nanotechonlogy	0.15	Desiccant	0.03
		Environmental	0.03	Magnetic	0.12	Pharmacy	0.02
				Metallosilicates	0.12		
				Metal powder	0.12		
				Fuel	0.12		
				Clay	0.12		
				Surface	0.12		
				Inorgan Chem	0.09		
				Membrane	0.09		
				Titanium	0.09		
				Thermochim	0.09		
				Kinetics	0.09		
				Solid Chem	0.06		
Median: 0.03, STDEV: 0.07, SKEW: 3.09		Median: 0.03, STDEV: 0.04, SKEW: 2.31		Median: 0.09, STDEV: 0.18, SKEW: 2.32			

citation types, we calculated science linkage (SL) for each, respectively, using the following formula by Computer Horizons Inc.:

$$SL = \frac{number.of.science.papers.cited.by.patents}{total.number.of.patents}$$

This resulted in three science linkage patterns, one for each citation type. The domains identified by these patterns are listed in order of largest to smallest contribution with their science linkage score in Table 3, by citation type. Pattern (a) of Table 3 shows the 13 scientific domains identified from the 176 scientific papers self-cited by inventor/applicant. Of these, Materials Science makes the largest contribution, followed by Electrical Chemistry, Physics Chemistry, Mesoporous Material Research, and Platinum Research. Pattern (b) of Table 3 shows the 25 domains identified from the 271 scientific papers cited by examiner. Of these, Electrical Chemistry makes the largest contribution followed by Materials Science, Zeolite Research, Chromatographic Research, and Physics. Pattern (c) of Table 3 shows the 61 scientific domains identified form the 2,205 non-self-cited scientific papers cited by the inventor/applicant. Of these, Physics Chemistry made the largest contribution followed by Materials Science, Physics, Electrical Chemistry, and Organic Chemistry.

Judgment by Domain Experts

All three sets of domains and science linkage calculations, together with an evaluation questionnaire, (see Figure 1) were sent to 267 experts in the domain of catalyst. The survey audience list came from the intersection of a *Science Citation Index* (SCI) author list and a DII inventor/applicant list; this enabled us to contact the domain experts with experience in paper publication and patent filing. Feedback from 39 counties (including China, U.S., France, Germany, Italy, Brazil,

Questionnaire for experimental results evaluation
Please choose one of the options after examining pattern (a):
Please choose one of the options after examining pattern (b):
Please choose one of the options after examining pattern (c):
Options:
Ⅰ all of the scientific domains are factually linked with the technology of catalyst Ⅱ most of the scientific domains are factually linked with the technology of catalyst Ⅲ majority of the scientific domains are factually linked with the technology of catalyst Ⅳ half the scientific domains are factually linked with the technology of catalyst Ⅴ minority of the scientific domains are factually linked with the technology of catalyst Ⅵ few of the scientific domains are factually linked with the technology of catalyst Ⅶ none of the scientific domains are factually linked with the technology of catalyst all=100% most=80%-100% majority=60%-80% half=40%-60% minority=20%-40% few=0%-20% none=0%

FIG. 1. Questionnaire provided to experts to report their evaluation of citation patterns.

TABLE 4. Distribution of cited papers within the IPC domains.

	Pattern (a): Inventor/applicant self-cited		Pattern (b): Examiner cited		Pattern (c): Inventor/applicant non-self-cited	
Citation number	176		271		2205	
Domains identified	13		24		61	
Experts' analysis of domains						
all domains factually linked	42.32%	83.14%	3.75%		0.75%	
most domains factually linked	22.47%		42.32%	76.77%	7.49%	
majority of domains factually linked	18.35%		21.72%		11.61%	
half of domains factually linked	7.87%		12.73%		9.36%	
minority of domains factually linked	4.87%		9.36%		37.08%	70.78%
few domains factually linked	1.50%		6.37%		16.85%	
no domains factually linked	0.37%		0.37%		13.48%	
none of the above	2.25%		3.37%		3.37%	
	100.00%		100.00%		100.00%	

Sweden, India, Russia, Japan, Korea, Spain, Mexico, Poland, New Zealand, Canada) was received from researchers in universities (e.g., Cornell University, U.S.; University of Stuttgart, Germany; University of Sussex, UK; University of Barcelona, Spain; Université de Nice, France; Tsinghua University, China), research institutes (e.g., Massachusetts Institute of Technology, U.S.; Korea Institute of Science and Technology; National Institute of Renewable Energy, India; Russia Urals Electrophysics Institute; Bayer Materials Science Institute, Germany; Shanghai Institute of Organic Chemistry, China) and innovators in enterprises (e.g., Shell Oil Co., Tokyo Gas Co. Ltd., China Petrochemical Co. Ltd., Ford Motor Co., Mitsubishi Chemistry Corp., Germany Global Technology Operations Inc.).

Expert review of the three domain sets is displayed in Table 4. The majority of experts (83.14%) believed that *all*, *most,* or a *majority* of the scientific domains were factually linked with the technology of catalyst when analyzing the 13 domains identified by the self-cited papers of inventor/applicants. When analyzing the 24 domains identified by examiner citations, a majority of experts (76.77%) believed that *mos*t, a *majority*, or *half* of the scientific domains were factually linked with the technology of catalyst. Conversely, after examining the 61 domains identified by inventor/applicants' non-self-citations, a majority of experts (70.78%) believed that a *minority*, *few*, or *no* scientific domains were factually linked with the technology of catalyst.

DOI: 10.1002/asi

Hypothesis Test

The relatively positive evaluation of most of the domain patterns ranked by science linkage within the citation type sets proves our H1 hypothesis: *Scientific papers cited in patents can be considered as a measure of science linkage*. The extremely positive evaluation (83.14%) of the self-cited by inventor/applicant citation pattern proves our H2a hypothesis: *Scientific papers self-cited by inventor/applicant are the best measure of linkage between science and technology*. The relatively positive evaluation of examiner citation pattern proves our H2b hypothesis: *Scientific papers cited by examiner rank second in indication of linkage between science and technology*. Finally, the mostly negative evaluation of the inventor/applicant non-self-cited citation pattern proves our H2c hypothesis: *Non-self-cited scientific papers by inventor/applicant are noisy and rarely indicate a linkage between science and technology*.

Discussion

The three pattern sets demonstrate a science linkage based on expert evaluation of a strong knowledge link between catalyst technology and the identified research domains. In Table 1 the top five scientific domains, strongly associated with catalyst technology, were similar, including Materials Science, Electrical Chemistry, Physics, and Physics Chemistry; however, each pattern shows individual characteristics.

Pattern (c) identifies 61 scientific domains linked to catalyst technology, only a minority of which were acknowledged as "factually linked with the technology of catalyst" by the experts. This indicates that non-self-citations by inventor/applicant are suffused with noise. It is likely that if we could isolate inventor/applicant non-self-citations that *describe theory background* and *explain knowledge source* alone, these could show science linkage. However, including inventor/application citations that *identify problems in practice* and *illustrate prior art defects or disadvantages* leads to indirect and even nonexistent science linkage. Hence pattern (c), based on non-self-cited scientific papers cited by the inventor/applicant, does not show science linkage objectively.

Pattern (a) identifies 13 scientific domains linked to catalyst technology. As more than 80% of experts acknowledge a majority of these to be "factually linked with the technology of catalyst," it is likely self-citations by an inventor/applicant can show science linkage precisely. However, as only a small portion of all inventors/applicants play the dual role of technical innovator and scientific researcher, only a few inventors/applicants have the ability to cite their own scientific paper in their patents. As a result, the quantity of self-citations by inventors/applicants is very small and the science linkage shown by them is only partial. Hence pattern (a), based on self-cited scientific papers cited by the inventor/applicant, can show science linkage precisely but incompletely.

Pattern (b) identifies 24 scientific domains linked to catalyst technology. As most, or at least half, of these are acknowledged as "factually linked with the technology of catalyst" by more than 80% of the experts, it is likely that citations by examiner can show science linkage accurately. In view of patent law, scientific papers cited by the examiner cite comparative prior art, which therefore carries the superordinate concept for comparison; thus the citing patent and the cited paper are connected logically in content. Furthermore, regardless of the inventor/applicant's self-cite papers, the examiner will cite the relevant literature for patentability examination. As the scientific domains indicated by examiner citation are larger in quantity than those self-cited by inventor/applicants, pattern (b), based on scientific papers cited by examiners, can show science linkage more accurately than non-self-citations by inventor/applicants, and more comprehensively than those self-cited by inventor/applicants.

Conclusion

We conducted an exploratory science linkage calculation in the technical domain "catalyst", based on three types of citations: self-citation by inventor/applicant, non-self-citation by inventor/applicant, and citation by examiner. According to the results, based on the evaluation of domain experts, we conclude that the non-self-citation by inventor/applicant citations are quite noisy and cannot indicate science linkage objectively; that self-citation by inventor/applicant citations can measure science linkage precisely but incompletely; and that examiner citations can indicate science linkage more accurately than inventor/applicant non-self-citations and more comprehensively than inventor/applicant self-citation.

A limitation of this study, however, is our use of only USPTO patents. Although they provided the largest data set with which to test our hypotheses, the differences between the USPTO and EPO requirements may mean our results are not transferable to the EPO system. Within the EPO system the search for prior art is carried out by the examiner, whereas in the USPTO system this is the burden of the inventor/applicant. Similar testing of the EPO system using our methods and hypotheses is a future challenge for our science linkage investigation.

Based on our findings, limitations notwithstanding, we suggest the following measures. When calculating the general relationship between science and technology where a complete view of linkage is required, all types of non-patent citations should then be used for measuring science linkage. However, citations self-cited by an inventor/applicant should be weighted highest, closely followed by examiner citations, and non-self-cited inventor/applicant citations should be weighted lowest. When calculating science–technology integration precisely where crucial knowledge linkage between science research and technology innovation is required, only self-cited by inventor/applicant citations should be used for measuring science

DOI: 10.1002/asi

linkage. In this case, examiner citations and non-self-citation by inventor/applicant citations should be excluded due to excessive noise.

Acknowledgments

This manuscript is based upon work supported by the funding initiative Discovery of Science-Technology Linkage. Specifically, funding comes from the Social Sciences and Humanities Research Fund of Ministry of Education, China (Grant No. 12YJC870013) and the National Social Science Fund, China (Grant No. 13BTQ056).

References

Alcácer, J., & Gittelman, M. (2006). Patent citations as a measure of knowledge flows: The influence of examiner citations. The Review of Economics and Statistics, 88(4), 774–779. doi:10.1162/rest.88.4.774

Alcácer, J., Gittelman, M., & Sampat, B. (2009). Applicant and examiner citations in U.S. patents: An overview and analysis. Research Policy, 38(2), 415–427. doi:10.1016/j.respol.2008.12.001

Allison, J.R., & Lemley, M.A. (1998). Empirical evidence on the validity of litigated patents. American Intellectual Property Law Association (AIPLA) Quarterly Journal, 26, 185. doi:10.2139/ssrn.118149

Azagra-Caro, J.M., Fernández-de-Lucio, I., Perruchas, F., & Mattsson, P. (2009). What do patent examiner inserted citations indicate for a region with low absorptive capacity? Scientometrics, 80(2), 441–455. doi:10.1007/s11192-008-2081-0

Azagra-Caro, J., Mattsson, P., & Perruchas, F. (2011). Smoothing the lies: The distinctive effects of patent characteristics on examiner and applicant citations. Journal of the American Society for Information Science and Technology, 62(9), 1727–1740. doi:10.1002/asi.21574

Bacchiocchi, E., & Montobbio, F. (2009). Knowledge diffusion from university and public research. A comparison between US, Japan and Europe using patent citations. The Journal of Technology Transfer, 34(2), 169–181. doi:10.1007/s10961-007-9070-y

Balconi, M., Breschi, S., & Lissoni, F. (2004). Networks of inventors and the role of academia: An exploration of Italian patent data. Research Policy, 33(1), 127–145. doi:10.1016/S0048-7333(03)00108-2

Bhattacharya, S., Kretschmer, H., & Meyer, M. (2003). Characterizing intellectual spaces between science and technology. Scientometrics, 58(2), 369–390. doi:10.1023/A:1026244828759

Breschi, S., & Catalini, C. (2010). Tracing the links between science and technology: An exploratory analysis of scientists' and inventors' networks. Research Policy, 39(1), 14–26. doi:10.1016/j.respol.2009.11.004

Breschi, S., Lissoni, F., & Montobbio, F. (2007). The scientific productivity of academic inventors: New evidence from Italian data. Economics of Innovation and New Technology, 16(2), 101–118. doi:10.1080/10438590600982830

Caballero, R.J., & Jaffe, A.B. (1993). How high are the giants' shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. NBER Macroeconomics Annual, 8(January), 15–74. doi:10.2307/3585017

Callaert, J., Van Looy, B., Verbeek, A., Debackere, K., & Thijs, B. (2006). Traces of prior art: An analysis of non-patent references found in patent documents. Scientometrics, 69(1), 3–20. doi:10.2139/ssrn.874906

Carpenter, M.P., & Narin, F. (1983). Validation study: Patent citations as indicators of science and foreign dependence. World Patent Information, 5(3), 180–185. doi:10.1016/0172-2190(83)90139-4

Chen, C., & Hicks, D. (2004). Tracing knowledge diffusion. Scientometrics, 59(2), 199–211. doi:10.1023/B:SCIE.0000018528.59913.48

Collins, P., & Wyatt, S. (1988). Citations in patents to the basic research literature. Research Policy, 17(2), 65–74. doi:10.1016/0048-7333(88)90022-4

Criscuolo, P., & Verspagen, B. (2008). Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. Research Policy, 37(10), 1892–1908. doi:10.1016/j.respol.2008.07.011

Grupp, H. (1996). Spillover effects and the science base of innovations reconsidered: An empirical approach. Journal of Evolutionary Economics, 6, 175–197. doi:10.1007/BF01202593

Hall, B. (2000). Innovation and market value. In R. Barrell & G. Mason (Eds.), Productivity, innovation and economic performance (pp. 177–198). Cambridge, UK: Cambridge University Press.

Hall, B.H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. Rand Journal of Economics, 36(1), 16–38. doi:10.1057/jors.1995.38

Harhoff, D., Scherer, F.M., & Vopel, K. (2003). Citations, family size, opposition and the value of patent rights. Research Policy, 32(8), 1343–1363. doi:10.1016/S0048-7333(02)00124-5

Hu, D., Chen, H., Huang, Z., & Roco, M.C. (2007). Longitudinal study on patent citations to academic research articles in nanotechnology (1976–2004). Journal of Nanoparticle Research, 9(4), 529–542. doi:10.1007/s11051-007-9215-9

Jaffe, A.B., Fogarty, M.S., & Banks, B.A. (1998). Evidence from patents and patent citations on the impact of NASA and other federal labs on commerical innovation. The Journal of Industrial Economics, XLVI(2), 183–205.

Jaffe, A.B., Trajtenberg, M., & Fogarty, M.S. (2000). Knowledge spillovers and patent citations: Evidence from a survey of inventors. American Economic Review, 90(2), 215–218. Retrieved from http://www.jstor.org/stable/117223

Jaffe, A.B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics, 108(3), 577–598. doi:10.2307/2118401

Karvonen, M., & Kässi, T. (2013). Patent citations as a tool for analysing the early stages of convergence. Technological Forecasting and Social Change, 80(6), 1094–1107. doi:10.1016/j.techfore.2012.05.006

Kesan, J.P. (2002). Carrots and sticks to create a better patent system. Berkeley Technology Law Journal, 17(2), 763–797. doi:10.2139/ssrn.305999

Lai, K., & Wu, S. (2005). Using the patent co-citation approach to establish a new patent classification system. Information Processing & Management, 41(2), 313–330. doi:10.1016/j.ipm.2003.11.004

Lampe, R. (2012). Strategic citation. The Review of Economics and Statistics, 94(February), 320–333. doi:10.1162/REST_a_00159

Meyer, M. (2000). What is special about patent citations? Differences between scientific and patent citations. Scientometrics, 49(1), 93–123. doi:10.1023/A:1005613325648

Narin, F., Hamilton, K.S., Olivastro, D. (1995). Linkage between agency-supported research and patented industrial technology. Research Evaluation, 5(3), 183–187. doi:10.1093/rev/5.3.183

Narin, F., Hamilton, K.S., & Olivastro, D. (1997). The increasing linkage between U.S. technology and public science. Research Policy, 26(3), 317–330. doi:10.1241/johokanri.40.1080

Narin, F., & Noma, E. (1985). Is technology becoming science? Scientometrics, 6, 369–381.

Owen-Smith, J., & Powell, W.W. (2003). The expanding role of university patenting in the life sciences: Assessing the importance of experience and connectivity. Research Policy, 32(9), 1695–1711. doi:10.1016/S0048-7333(03)00045-3

Rip, A. (1992). Science and technology as dancing partners. In P. Kroes & M. Bakker (Eds.), Technological development and science in the Industrial Age (Vol. 144, pp. 231–270). Dordrecht: Springer Netherlands. doi:10.1007/978-94-015-8010-6

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. doi:10.1016/j.respol.2006.09.014

Schmoch, U. (1993). Tracing the knowledge transfer from science to technology as reflected in patent indicators. Scientometrics, 26(1), 193–211. doi:10.1007/BF02016800

Schumpeter, J.A. (1942). Capitalism, Socialism and Democracy. New York: Harper Brothers.
Sternitzke, C. (2009). Reducing uncertainty in the patent application procedure—Insights from invalidating prior art in European patent applications. World Patent Information, 31(1), 48–53. doi:10.1016/j.wpi.2008.04.007
Stock, M., & Stock, W.G. (2006). Intellectual property information: A comparative analysis of main information providers. Journal of the American Society for Information Science and Technology, 57(13), 1794–1803. doi:10.1002/asi
Thompson, P. (2006). Patent citations and the geography of knowledge spillovers: Evidence from inventor- and examiner-added citations. The Review of Economics and Statistics, 88(2), 383–388. doi:10.1162/rest.88.2.383
ThomsonReuters. (2011). 2011 Journal Citation Reports.
Tijssen, R.J. (2001). Global and domestic utilization of industrial relevant science: Patent citation analysis of science–technology interactions and knowledge flows. Research Policy, 30(1), 35–54. doi:10.1016/S0048-7333(99)00080-3
Tijssen, R.J.W. (2002). Science dependence of technologies:Evidence from inventions and their inventors. Research Policy, 31(4), 509–526. doi:10.1016/S0048-7333(01)00124-X
Verbeek, A., Debackere, K., & Luwel, M. (2003). Science cited in patents: A geographic "flow" analysis. Scientometrics, 58(2), 241–263. doi:10.1023/A:1026232526034
Weinstock, M. (1970). Citation indexes. In A. Kent & H. Lancour (Eds.), Encyclopedia of library and information science (pp. 16–40). New York: Marcel Dekker, Inc.
Wong, P.-K., & Ho, Y.-P. (2007). Knowledge sources of innovation in a small open economy: The case of Singapore. Scientometrics, 70(2), 223–249. doi:10.1007/s11192-007-0201-x