Journal of Informetrics 13 (2019) 100985

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Exploring the forward citation patterns of patents based on the evolution of technology fields

Su Jung Jee, Minji Kwon, Jung Moon Ha, So Young Sohn *

Department of Industrial Engineering, Yonsei University, 134 Shinchon-dong, Seoul 120-749, Republic of Korea

ARTICLE INFO

Article history:
Received 6 April 2019
Received in revised form 20 October 2019
Accepted 28 October 2019
Available online 10 November 2019

Keywords:
Forward citation patterns of patents
Dynamic field effects
Technology diffusion
Technology value
Evolutionary perspective

ABSTRACT

Forward citations of patents have been used extensively to capture the impact of technological knowledge. However, our understanding of the factors shaping patent citation patterns remains limited. One of the main limitations is the lack of scholarly attention paid to the dynamic influences arising from the evolution of technology fields. From an evolutionary perspective, technological impact is not simply determined by the static attributes of a technology itself; it is also dynamically affected by changes in the external conditions. Drawing on this viewpoint, this study suggests a model for understanding patent citation patterns by reflecting the evolution of the technology fields to which each patent belongs. Four such factors are explored: technology cycle time, potential of technological convergence, popularity of the technology field, and technological novelty. Based on the proposed model, we show how expected citation patterns can change as a result of different scenarios for technology field evolution. We conduct a case study of patents in the information technology and healthcare industries to show citation patterns of patents across heterogeneous industries as well as those within an industry. Contributions to the innovation literature and research investment decisions are discussed.

1. Introduction

Forward citations of patents have long been used to identify and analyze the diffusion and value of technological knowledge. Jaffe and Trajtenberg (1996) developed a citation function that improved our understanding of how technological knowledge spills over across geographic and institutional boundaries over time. Their work inspired numerous scholars to use forward citation information to capture the flow of technological knowledge in a variety of contexts. In addition, prior studies have validated that counts of forward citations of patents can be used as a proxy for the importance and market value of a technology (Albert, Avery, Narin, & McAllister, 1991; Hall, Jaffe, & Trajtenberg, 2005; Harhoff, Narin, Scherer, & Vopel, 1999; Trajtenberg, 1990). The empirical evidence provided by such scholars has resulted in a number of subsequent studies; these have assumed that citation counts can be used to represent the value of a technology. That is, forward citations of patents have been widely used by both academic researchers and practitioners interested in measuring the impact of technological knowledge.

Together with this broad employment of forward citation information, there has been increased demand to reveal the factors shaping patent citations, as such insights can guide the evaluation of current technologies as well as decision making

* Corresponding author.
E-mail addresses: jsj930jsj930@hanmail.net (S.J. Jee), kminji@naver.com (M. Kwon), jungmoon27@hanmail.net (J.M. Ha), sohns@yonsei.ac.kr (S.Y. Sohn).

https://doi.org/10.1016/j.joi.2019.100985

about investment in research and development (R&D). However, our understanding of the factors influencing the patterns of patent citations is still limited. One of the reasons is that prior studies have primarily focused on the relationship between the static attributes relevant to patents and the citations the patents receive (Albert et al., 1991; Carpenter, Narin, & Woolf, 1981; Jaffe & Trajtenberg, 1996; Reitzig, 2003), while relatively ignoring the myriad of dynamic influences of the technology fields to which the focal patent belongs.

Although the static method of investigation used by scholars has revealed some crucial insights into the factors influencing patent citations, this approach has been unable to incorporate the time-varying factors that shape citation patterns as patents age. Moreover, how such dynamic influences differ across technology fields remains a question. Jaffe and De Rassenfosse (2017) also highlighted that the extent to which time and technology field affect patents' forward citations is understudied. For example, the following aspects have rarely been investigated: which technology field conditions lead some patents to be cited more or less frequently than others by a certain patent age, which technology field conditions influence the earlier or later appearance of the peak of expected probability for the next citation of a patent, and how such dynamic influences from technology field evolution differ by the field to which a patent belongs. Providing answers to those questions may help decision makers interested in the expected impact of certain technologies.

This gap in the literature aligns with the evolutionary view of technological change, which is one of the most widely accepted viewpoints for understanding the impact of technologies. This perspective highlights the dynamic nature of technology that continuously evolves and, in particular, maintains that the technology selection mechanism does not simply depend on the innate attributes of a technology itself but is dynamically linked to the complex environmental conditions with which the technology is associated (Dosi, 1982; Nelson & Winter, 1982). Following this viewpoint, this study aims to bridge the gap in the literature by investigating how the evolution of the technology field dynamically influences the citation patterns of a patent as it ages. Although the environmental conditions noted by evolutionary scholars include the social, legal, and political dimensions as well as the technology field, this study focuses on the technological aspect.

Our exploration reveals how different evolutionary scenarios at the technology field level result in different patterns in the expected probability of the next citation over the various phases of patent age. The proposed model considers the evolution of multiple sub-technology fields associated with each patent; these fields can lie either across industries or within an industry. We employ a Cox model, which accommodates not only the dynamic environmental features related to a focal patent but also the right-censored characteristic of forward citation data. Our model focuses on the four dimensions of time-varying factors expected to affect patent citation patterns: technology cycle time (TCT), potential of technological convergence, popularity of the technology field, and technological novelty. Moreover, we control for possible patent-level attributes, which are static but known to affect forward citations.

The proposed model is applied to 20 years (1995—2014) of United States Patent and Trademark Office patents related to the information technology (IT) and healthcare industries. Our investigation shows how different scenarios for the four aforementioned environmental dimensions result in changes in the expected forward citation patterns of patents and how such changes differ for patents applied across industries than within an industry. This study is expected to enhance the fundamental understanding of the use of citation information and contribute to R&D investment decision making for both industry and governments.

The remainder of this paper is composed of the following sections. Section 2 reviews the literature on the use of patent citation information, the evolutionary perspective of technological change, modeling citation information, and technological convergence. Section 3 introduces the key factors of our model and Section 4 shows the results of its application. Lastly, Section 5 discusses the academic and practical contributions of the study along with its limitations.

2. Literature review

2.1. Use of patent citation information

A patented invention should disclose any novel aspect usually generated from the recombination of previous knowledge (Gilfillan, 1935; Nelson & Winter, 1982). In a patent document, the reference section reveals the prior knowledge used or recombined to develop the new invention. Hence, earlier patents cited in a new patent document can be understood as the pieces of knowledge that influence the creation of a new invention. Relying on the nature of the relationship between the cited and citing patents, a number of studies have used citation information to analyze two main areas: the estimation of patent value in terms of its influence (Albert et al., 1991; Hall et al., 2005; Harhoff et al., 1999; Trajtenberg, 1990) and the diffusion of knowledge within and across institutional, geographical, and even cultural boundaries (Alcacer & Gittelman, 2006; Fung & Chow, 2002; Jaffe & Trajtenberg, 1996; Jaffe, Trajtenberg, & Henderson, 1993; Karvonen & Kässi, 2013; Thompson, 2006).

First, studies employing citation information to estimate the value of a technology have used the number of forward citations during a unit period as a proxy for the impact or performance of a technology. Such employment is roughly based on Trajtenberg (1990), who examined the significant relationship between the number of forward patent citations and economic performance, captured from the utilization of the patent. Other studies provide detailed empirical evidence to support the use of citation counts as a proxy for patent value (Albert et al., 1991; Hall et al., 2005; Harhoff et al., 1999; Lanjouw & Schankerman, 1999). This validation of the use of citation information has made a considerable contribution to the innovation literature by providing a more objective tool to represent the economic value of technology, although this method of measurement is still incomplete. A number of subsequent studies have revealed the factors resulting in higher innovation performance at various levels, including the firm (e.g., Huang & Chen, 2010), region (e.g., Mukherji & Silberman, 2013), and even country (e.g., Hu & Mathews, 2008).

Second, citation information has also been widely used to understand the patterns of technological knowledge diffusion. Jaffe and Trajtenberg (1996) proposed a citation function to investigate technological knowledge spillovers over time. Many subsequent scholars have investigated knowledge spillovers with respect to other crucial dimensions by relying on patent citation information. For instance, Branstetter (2001) explored the role of intra- and international knowledge spillovers in innovation and productivity in Japan and the United States, and Fung and Chow (2002) showed how technological knowledge diffuses both within an industry and among industries. In addition, Stolpe (2002) explored the determinants of knowledge diffusion by focusing on changes in inventors' affiliations, and Alcacer and Gittelman (2006) and Thompson (2006) used patent citation information to investigate the different impacts of examiner and inventor citations on knowledge spillovers. Although some scholars have mentioned the incompleteness of patent citations to reflect technological knowledge spillovers (Jaffe, Trajtenberg, & Fogarty, 2000), recent evidence shows that technological content similarities between cited and citing patents are much higher than those of non-citing/cited pairs of patents (Chen, 2017).

Along with this broad use of patent citation information, the demand to understand the factors that lead some patents to be more heavily cited has increased in both academia and industry. This is because such an insight could help not only anticipate the impact of current technologies but also make R&D investment-related decisions. However, although numerous studies have explored the factors relevant to patent citations, most previous approaches have overlooked the dynamic and continuous influences coming from the technological environment surrounding each patent (Jaffe & De Rassenfosse, 2017). Rather, they have usually focused on the predetermined factors related to patents. This static approach may mislead expectations about the impact of technologies because changes in environmental circumstances may result in different patent citation patterns from those expected by simply considering static influences. This study bridges the gap in the literature by proposing a model incorporating various dimensions of the technological environment surrounding a focal patent and by investigating how different scenarios of environmental conditions change patent citation patterns.

2.2. Evolutionary perspective of technological change

The gap we discuss above aligns with the basic tenet of the evolutionary view of technological change. This viewpoint draws an analogy between the biological evolutionary process and technological evolutionary process, which are similar in terms of the generation of numerous variations, competition among alternatives, and consequences for winners and losers (Nelson & Winter, 1982). Nowadays, scholars in diverse fields of study addressing technology acknowledge that the process of technological change is akin to the process of biological evolution (Nelson et al., 2018).

The most fruitful outcome of the evolutionary view of technological change is the concept of a "technological paradigm," which is an abstract notion that encompasses all the factors influencing the evolution of a technology (Dosi, 1982). The technological paradigm at a specific timepoint is described as the body of technological artifacts and processes, the way they are used, the body of understanding about a technology, and an assessment of the prevailing technical options, to name a few. The concept is similar to the scientific paradigm, defined by Kuhn (1963), and technological regime, defined by Nelson and Winter (1982). Such perspectives of technology commonly imply that a technology is not isolated or static; rather, it should be understood as existing within a dynamically evolving system. Through the evolutionary process, technologies representing high fitness to their environment receive more attention than other technologies that exhibit a lower degree of fitness. The environment surrounding a technology has been described as a collection of other technologies, market conditions including customers' needs, and even regulations and political circumstances. However, prior studies exploring the factors influencing patent citations have been insufficient to incorporate this viewpoint into their investigations (Albert et al., 1991; Carpenter et al., 1981; Jaffe & Trajtenberg, 1996; Reitzig, 2003). The present study relies on this viewpoint to explore the forward citation patterns of patents, with a particular focus on the evolving technological environment with which each focal patent is associated.

2.3. Modeling citation information

Studies that have investigated the factors affecting patent citations have modeled citation data in various ways. Jaffe and Trajtenberg (1996) modeled the process of the generation of patent citations using a double-exponential function, which has the advantage of accommodating both the diffusion and the obsolescence of a technology simultaneously. Although this model provides a useful tool to investigate knowledge spillovers with respect to patent age, this model is limited because of its neglect of the huge proportion of patents that have never been cited. Moreover, although every citation event is right censored, because we can only obtain events up to the point of our observation, the right-censored character of citation information has not been considered in this model. Another approach for modeling citation data relies on the Poisson or negative binomial model, which uses citation counts to find the determinants arousing further citations (Hausman, Hall, & Griliches, 1984). Since a large proportion of patents are never cited during their lifetimes, scholars have also used a zero-inflated Poisson or zero-inflated negative binomial model to reflect this circumstance (Lee, Lee, Song, & Lee, 2007). Despite this benefit, such models still have limitations since they also fail to consider the right-censored nature of citation information; further, these models are too simple to capture any time-varying influences.

Marco (2007) supplemented the censoring problem by employing hazard estimation based on the inter-citation time. This model considers not only the right-censoring issue but also makes it possible to estimate the hazard rate of subsequent citations over time. However, the research focus of Marco (2007) was showing that heterogeneity exists in citation dynamics rather than investigating the factors influencing citation rates. Therefore, although the author's approach has merits in terms of accommodating censored information and heterogeneity in individual patents, environmental conditions that vary over time need to be further investigated. To incorporate the time-varying influences of technological field-level evolution to estimate citation dynamics and experiment with various scenarios of changing environmental conditions, we adopt Marco's (2007) inter-citation time model, which employs the hazard rate estimation. Using this model, we can take advantage of three major aspects. First, we can perform a longitudinal analysis for the forward citation history of each patent. Second, we can accommodate the censored nature of citation information and even include patents with zero citations. Finally, and most importantly, we can reflect the evolution in the relevant technological fields faced by each patent during its lifetime.

2.4. Technological convergence

The importance of our approach is underlined by the increasing trend of technological convergence in many industries. This is particularly true in the rise of information and communications technologies (ICT), in which technology has increasingly been created by combining knowledge in new and unexpected ways. The blurred boundaries among technological fields have spurred many scholars to study and explain the growing complexity underlying the phenomenon. Indeed, before the emergence of the ICT industry, Rosenberg (1963) had already mentioned the concept of technological convergence in the context of machine tools and metal-related industries to indicate the process by which the technological bases supporting different industries gradually overlap. After the emergence of the ICT industry, a number of scholars actively adopted and used the concept of technological convergence to characterize the vigorous trend by which several industries employ ICT (Curran & Leker, 2011; Karvonen & Kässi, 2013). Because of the complex nature underlying such a phenomenon, many concerns have been aroused around how to estimate and understand the value, lifespan, and spillover impact of convergence technologies (Preschitschek, Niemann, Leker, & Moehrle, 2013). Practically, this is an important issue from the perspective of R&D investment decision making in many industries. We expect our approach to be a useful decision making tool for such people.

The range of concepts connoted by the term "technological convergence" is broad when we consider the variation in the degree of similarity among several converging areas. In other words, some technologies are created by multiple converging sub-fields within a specific industry, while others can be created from the combination of fundamentally heterogeneous fields. To distinguish the type of convergence by the degree of variation among converging domains, scholars have proposed the concept of within- and inter-industry convergence (Hacklin, 2007; Kim, Lee, Kim, Lee, & Suh, 2015; Van den Berghe & Verweire, 2001). Within-industry convergence occurred exceptionally frequently in the ICT industry during the high-technology revolution beginning in the 1970s (Davies, 1996; Duysters & Hagedoorn, 1998). In addition, within-industry convergence has occurred around traditional industries as well as high-tech electronics-based industries. For instance, in the financial service industry, convergence between business model technologies related to banking and insurance and to insurance and pension markets has occurred (Van den Berghe & Verweire, 2001). By contrast, numerous convergences are occurring among different industries (Hacklin, 2007; Kim et al., 2015). A typical example mentioned is the convergence between the IT and biotechnology industries (Geum, Kim, Lee, & Kim, 2012). Recently, electric vehicles are the final products resulting from technological convergence between the traditional automobile and chemical industries (von Delft, 2013). Because patterns of patent citations may differ for the two groups of patents (i.e., patents within an industry and patents across industries), our approach is designed in a way to compare both circumstances.

3. Data and methodology

3.1. Data

We use patent applications to the United States Patent and Trademark Office (USPTO) between 1995 and 2014. To collect patent information, we use the Worldwide Patent Statistical Database, which provides patent information collected from international patent offices including the USPTO. Our analysis uses both granted and non-granted patents because forward citation events occur in both groups of patents. In addition, to reflect the patents involved in one or several industries, this study chooses two heterogeneous industries, healthcare and IT. Although they differ, they have increasingly interacted with each other in recent decades (Touati & Tabish, 2013). Choosing these two industries thus allows us to compare the citation patterns of patents within an industry as well as across industries.

To collect patents related to these two industries, we rely on International Patent Classification (IPC) information, which is a hierarchical patent classification system used as a primary way to search for patents by technology field (WIPO, 2019). Based on the IPC classes (i.e., three-digit IPC) related to the healthcare and IT industries (see Table 1), we collect patents that belong to at least two of the eight classes listed, namely G05, G06, G07, G08, G09, A23, A61, and B01. Out of these collected patents, we exclude those which have additional classes that are not listed in Table 1. By using patents that belong to at least two classes listed in Table 1, we focus on investigating the citation patterns of patents related to convergence both within an industry (healthcare or IT) and across these two heterogeneous industries (healthcare and IT). In addition, by excluding

Table 1
Industries and technology classes addressed.

Industry	Technology Class	Description
IT	G05	Controlling; Regulating
	G06	Computing; Calculating; Counting
	G07	Checking-devices
	G08	Signaling
	G09	Educating; Cryptography; Display; Advertising; Seals
Healthcare	A23	Foods or food stuffs; Their treatment, not covered by other classes
	A61	Medical or veterinary science; Hygiene
	B01	Physical or chemical processes or apparatus in general

patents that have additional classes not listed in Table 1, we can focus on the influence of the technology fields related to our target industries. As a result, we obtain 65,960 US patents.

Then, for the collected patents, we checked the forward citations they received until the end of 2014. Out of the 65,960 patents, 14,373 patents did not receive any forward citations, while the remaining patents received at least one forward citation from their application date to the end of 2014. Our model reflects the censored time for every patent as the last day of 2014. In addition, each forward citation is considered to be an individual event for modeling the inter-citation time. In total, there are 889,250 records covering both censored and citation events. Because there are multiple observations for each focal patent, we use robust standard errors clustered by cited patent in the estimation.

3.2. Covariates reflecting technology field evolution

Each forward patent citation event is accompanied by time-varying covariates, which reflect the evolutionary state of the technology fields in which the focal patent is embedded. This study considers the state of the following four dimensions: TCT, potential of technological convergence, popularity of the technology field, and technological novelty. This section provides the rationale behind the choice of these dimensions and how they are measured.

3.2.1. Technology cycle time

Haupt, Kloyer, and Lange (2007) showed that the speed of technological advancement is significantly higher when a technology is in the growth stage than in the introduction or maturity stages. In other words, the rate of technological progress may differ depending on the stage in the technology life cycle. The rate of progress for a technology field is closely related to the number of patents applied for in that area. Therefore, one can expect the number of citations a patent receives to be higher when it is involved in a more rapidly progressing area. At the same time, technologies tend to be quickly replaced by subsequent technologies when the field is rapidly progressing (Haupt et al., 2007). In this vein, one can also expect the number of citations a patent receives to be lower when the patent belongs to a more rapidly advancing technology field. Therefore, both positive and negative effects of the rate of technological progress on the citation rate seem possible. We can at least expect that the rate of technological progress may dynamically influence the citation pattern of a focal patent.

The rate of technological progress has usually been determined through the citation lag between prior and future technologies, which is termed TCT (Kayal & Waters, 1999). The widely used operational definition of TCT is the median age of the patents cited in a patent, capturing the immediacy of reliance on prior knowledge. A lower TCT value implies that the technology is developing at a rapid pace by relying on relatively recent technologies. On the contrary, a higher TCT value implies that the speed of technological progress is relatively slow and that the technology is based on older technologies (Narin, 1993). Since we are interested in taking field-level progress into account, we aggregate the TCT information by patent class. To obtain the TCT for each class in year T, we compute the mean TCT for patents within each class over the preceding five years (from T-4 to T), which is a tentative duration for determining the recent speed of technological progress at year T. As a result, we obtain TCT values for the eight classes included in the IT and healthcare industries for each year from 1995 to 2014. Finally, to obtain the rate of field-level technological progress related to each patent at year T, the calculated TCT values of the fields related to each patent are averaged. A higher value of averaged TCT indicates a slower rate of technological progress in the fields relevant to each patent.

3.2.2. Potential of technological convergence

If a specific technology field tends to frequently interact with other technology fields, we can expect the likelihood that the focal field would experience a technological convergence with other areas to be high since the range of technologies interacting with the focal field is broad. Such fields are likely to gain attention from actors in various domains, and attention from other areas is closely related to the number of citations a patent in that field receives. Therefore, we postulate that the convergence potential of the technology fields in which a focal patent is embedded may influence the citation pattern of the focal patent. This is a time-varying factor because technological convergence repeatedly appears and disappears over time (Kim, Kim, & Lee, 2019).

To reflect the time-varying convergence potential of each technology field, we measure the degree centrality from an IPC co-occurrence network, which has been used to analyze technological convergence (Jee & Sohn, 2015). The co-occurrence of IPCs has been widely used to measure technological convergence (Curran & Leker, 2011; Karvonen & Kässi, 2011). In the IPC co-occurrence network, each node corresponds to each IPC and each link connects two IPCs if the classes have ever co-occurred within the same patent. A higher value for degree centrality in the IPC co-occurrence network implies that the class has been combined with many other fields. Based on this concept, we obtain the value for the convergence potential of each technology field in year *T* by calculating how broadly each class has co-occurred with other classes over the preceding five years (from *T-4* to *T*). The degree centrality of each field in the IT and healthcare industries is obtained for each year during 1995—2014. Then, to obtain the field-level convergence potential related to each patent at year *T*, the calculated centralities of the fields related to each patent are averaged.

3.2.3. Popularity of the technology field

The popularity of a technology field represents how many actors are interested in developing relevant technologies and capturing value from those technologies. The popularity of a specific technology field is also a time-varying attribute since attention on a specific domain changes over time because of various factors such as market demand (Dosi, 1982), the stage of the technology life cycle (Andersen, 1999), and the existence of complementary infrastructure (Nalebuff, Brandenburger, & Maulana, 1996) at a particular timepoint. The changing degree of attention on a technology field might be dynamically linked to the number of citations a patent in the field receives during its lifetime. Hence, we expect that the popularity of the technology field can be used as a key covariate to model citation patterns by the age of a patent.

Since it is meaningful to take into account the combination of knowledge used in a technical field, the popularity of the technology field is calculated for each pair of patent classes at year *T* by measuring the number of patents related to each class pair over the preceding five years (from *T-4* to *T*). Hence, the popularity of the technology fields in the IT and healthcare industries is obtained for each year from 1995 to 2014. Finally, to obtain the popularity of the technology fields related each patent at year *T*, we use the maximum value among the calculated popularity values related to each patent.

3.2.4. Technological novelty

Various definitions of technological novelty are introduced in the literature. This study draws on the definition based on knowledge recombination that has been widely accepted by innovation and management scholars. Although some scholars have argued that an original technology does not build on any prior knowledge (Ahuja & Morris Lampert, 2001; Banerjee & Cole, 2011), such truly original technologies are extremely rare (Strumsky & Lobo, 2015). Many scholars have concluded that novel invention ideas are likely to emerge from the recombination and refinement of diverse pieces of existing knowledge (Kogut & Zander, 1992; Strumsky & Lobo, 2015; Uzzi & Spiro, 2005).

Based on this concept, the literature has investigated whether novel technological ideas are more or less likely to be highly impactful knowledge, but the results from relevant studies have been inconsistent. Some scholars have noted that novelty in knowledge positively influences its impact (Ahuja & Morris Lampert, 2001; Kim, Cerigo, Jeong, & Youn, 2016; Simonton, 1999), while others have concluded that technological novelty is accompanied by higher risk, as proven by the higher dispersion of its impact (Verhoeven, Bakker, & Veugelers, 2016). Moreover, some scholars have even shown that there is no straightforward relationship between the novelty of an invention and its impact (Strumsky & Lobo, 2015). Although varied empirical evidence has been presented on the relationship between the novelty of technological knowledge and its influence, such evidence implies the possibility that the novelty of the knowledge underlying a patent might be related to citation patterns. The novelty of ideas in a patent is not a static characteristic but a dynamic one because the concept itself connotes a relative comparison with other technologies in a specific period.

To measure the dynamically changing degree of the technological novelty of a patent, we follow Uzzi, Mukherjee, Stringer, and Jones (2013), who suggested using a z-score to measure the novelty of the recombined knowledge of a research paper. The authors suggest measuring novelty by comparing the expected and observed frequency of knowledge recombination. Thus, a more negative z-score represents a higher novelty of combined knowledge. Kim et al. (2016) applied an identical concept to calculate the novelty of knowledge in terms of combined patent classes. We follow the method proposed by Kim et al. (2016) to calculate the novelty of a patent class combination, which is evaluated in year *T*. The z-score for the classes *p* and *q* in year *T* is defined as follows:

$$z_{pqT} = \frac{o_{pqT} - \mu_{pq}}{\sigma_{pq}},$$

where o_{pqT} is the observed number of patents belonging to *p* and *q* in year *T* and μ_{pq} and σ_{pq} are the expected number and standard deviation of patents that belong to *p* and *q* in the preceding five years (i.e., *T-5* to *T-1*). The z-scores of individual pairs of classes covering the IT and healthcare industry are obtained for each year from 1995 to 2014. To capture the ultimate novelty within each patent, we employ the minimum z-score among all the possible z-scores related to each patent.

3.3. Other patent-level covariates

In addition to the key environmental variables explained above, prior studies have examined several patent-specific variables that are static but can have a significant influence on the forward citations of patents. We control for such aspects, which are determined at the time the patent is filed and do not change over time. First, heterogeneity in R&D actors can lead to differences in R&D performance (Reagans & Zuckerman, 2001). To account for the effect of heterogeneous inventors engaging in a patent application, we control for the number of inventors (*number of inventors*). Previous studies have shown that the number of inventors has a positive and significant effect on the forward citations of a patent (Alnuaimi & George, 2016; He & Deng, 2007). Second, we use the number of assigned IPCs to control for the breadth of each patent (*number of IPCs*). A broader scope of a patent makes it harder for others to invent around the patent (Fischer & Leidinger, 2014). In line with this, Lerner (1994) found that the number of assigned IPCs is positively related to the value of the patent. Third, external research collaborations can provide R&D organizations with a way of creating synergy (Lavie & Rosenkopf, 2006). The actors who collaborate in an R&D project may be captured by the applicants of a patent (i.e., co-patentees) (*number of applicants*). Belderbos, Cassiman, Faems, Leten, and Van Looy (2014) found that co-patented inventions are more likely to be cited by future inventions, although self-citations are rather less likely for those patents. Fourth, we use the number of patent claims (*number of claims*), which is relevant to patent value (Tong & Frame, 1994). Claims represent the legal scope of a patent, and the broader the property rights protection, the more likely that others will rely upon the patent (Petruzzelli, Rotolo, & Albino, 2015). Lanjouw and Schankerman (1999) found that the number of claims in a patent is positively correlated with its forward citations. Fifth, we control for the number of references to prior patents (i.e., backward citations) (*number of references*). Patents that cite more prior arts are expected to be in technologically crowded areas and thereby have a different influence than other patents (Fleming, 2001). Some scholars have found a positive relation between the number of references in a patent and its forward citations (Nerkar & Paruchuri, 2005). Sixth, Mukherjee, Romero, Jones, and Uzzi (2017) found that patents that cite prior arts with a low mean age and high age variance are likely to be more highly cited than other patents. To consider this aspect, we control for the mean and variance of the ages of referenced patents in each focal patent in the application year (*references age mean, references age variance*). Lastly, we also use a dummy variable identifying whether the patent lies within a single industry (only IT or only healthcare) or across industries (both IT and healthcare) (*industry dummy*).

In addition to the time-invariant covariates, our model includes two other time-varying covariates at the patent level. First, we use the age of the focal patent when each forward citation event occurs (*age, updated monthly*). In addition, we include a squared term of patent age (*age squared*) to reflect the general expectation that patent citations have an inverse U-shaped pattern with age (Marco, 2007; Nerkar & Paruchuri, 2005; Podolny & Stuart, 1995). Second, we use the number of forward citations a patent has received divided by its age (*forward number/age*) to reflect the Matthew effect (i.e., patents that have already received more citations are more likely to be cited again).

3.4. Model

The dependent variable in our model is the time lag (day) between each sequential citation event. We use the semi-parametric Cox proportional hazard model:

$$\lambda_i(t) = \lambda_0(t) \exp\left(\beta z_i + \gamma x_i(t)\right),$$

where $\lambda_i(t)$ is the citation rate of patent *i* at time *t* (i.e., the subsequent citation event did not occur until *t*, but it occurs right after *t*); $\lambda_0(t)$ is a baseline citation rate that does not have any assumption regarding its distribution; and $x_i(t)$ indicates the time-varying covariates that reflect the field-level evolution with respect to the four dimensions (i.e., *TCT, potential of technological convergence, popularity of the technology field, technological novelty*), *age*, and *forward number/age*. Lastly, z_i is the notation for the time-invariant covariates for the innate and static properties of patent *i*.

4. Results and discussion

Table 2 shows the descriptive statistics and correlations of the variables used in our model and Table 3 shows the results of the Cox model. Model 1 in Table 3 is a baseline model with only the control variables. Models 2–5 sequentially add the technology field-level variables into the baseline model and Model 6 is the full model that includes all the variables. Compared with the baseline model (i.e., Model 1), Models 2–6 show a significant drop in the Bayesian information criterion (BIC) score (i.e., above 10 points), which implies that considering the field-level variables can provide a better explanation.

In line with prior studies (Nerkar & Paruchuri, 2005; Podolny & Stuart, 1995), *age squared* is negatively significant. This finding means that the rate of citation per patent generally increases as a patent ages but starts to decrease at a certain point. As expected from the Matthew effect, the number of forward citations a patent has received divided by its age (*forward number/age*) is positive and significant, implying that patents that have received more citations are more likely to receive another. Consistent with our expectations, *number of claims*, *number of references, number of inventors*, and *number of IPCs* are positive and significant in both the baseline and the full models. In addition, concurring with Mukherjee et al. (2017), *references age mean* and *references age variance* have negative and positive effects on the citation rate, respectively (both significant). On the contrary, *number of applicants* is negative and significant. This result can be partially explained by the fact that the relation between R&D collaboration and performance can differ by the collaboration conditions (Belderbos, Carree, & Lokshin, 2004). Lastly, the categorical variable used to indicate different groups of patents by industry (i.e., within IT versus within healthcare and across IT and healthcare versus within healthcare) is shown to be positively significant.

Model 2 adds *TCT* as a technology field-level variable, and this variable is shown to be positive and significant. This finding implies that the rate of technological progress, which is inversely proportional to the TCT, has a negative relationship with

Table 2
Descriptive statistics and correlations (n = 889,250).

Variable	Mean	SD	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. TCT (day)	2198	151.79	1906.75	2677.5	1.00													
2. Potential of technological convergence (degree centrality)	106.4	8.73	64.5	120.5	−0.20	1.00												
3. Popularity of technology field (number of patents)	6224	3176	2	12676	−0.19	0.32	1.00											
4. Technological novelty (minimum z-score)	−10	41	−98.103	72.315	−0.02	−0.55	−0.01	1.00										
5. Age (month)	75	48	0	243	0.20	−0.15	−0.13	−0.10	1.00									
6. Age squared	7860	9184	0	59049	0.20	−0.17	−0.15	−0.07	0.95	1.00								
7. Forward number/age	0.49	0.80	0	15	0.03	−0.08	0.08	0.09	0.12	0.11	1.00							
8. Number of inventors	2.77	2.12	1	29	0.01	0.05	0.01	−0.05	−0.04	−0.03	0.09	1.00						
9. Number of IPCs	2.05	0.23	2	5	0.02	−0.05	0.06	−0.20	0.00	0.00	0.00	0.03	1.00					
10. Number of applicants	1.8	1.78	0	30	0.07	0.04	0.03	−0.11	−0.20	−0.18	−0.05	0.43	−0.02	1.00				
11. Number of claims	22.12	22.43	0	466	−0.01	0.03	0.04	−0.01	0.08	0.07	0.18	0.06	0.04	−0.04	1.00			
12. Number of references	26.94	31.8	0	222	0.04	0.03	0.02	−0.05	−0.06	−0.06	0.12	0.14	0.05	0.18	0.20	1.00		
13. References age mean	85.43	49.6	0	1022.9	0.02	0.00	−0.21	−0.01	0.00	0.01	−0.04	0.02	0.04	0.02	0.02	0.30	1.00	
14. References age variance	4785.73	10075.4	0	427196.1	0.01	0.00	−0.16	0.02	0.01	0.01	−0.05	0.00	0.01	0.00	−0.01	0.11	0.62	1.00
15. Industry	Industry dummy: within IT (1), across healthcare and IT (2), and within healthcare (reference group).																	

Table 3
Results of the Cox regression.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
TCT (day)		6.449e-04***				9.216e-04***
		(2.163e-05)				(2.212e-05)
Potential of technological convergence (degree centrality)			2.174e-02***			3.814e-02***
			(8.534e-04)			(7.816e-04)
Popularity of technology field (number of patents)				2.681e-05***		−3.427e-06*
				(2.048e-06)		(2.071e-06)
Technological novelty (minimum z-score)					5.126e-03***	9.578e-03***
					(3.046e-04)	(2.806e-04)
Age (month)	7.254e-04***	2.875e-04	4.636e-04**	5.394e-04**	2.068e-03***	1.964e-03***
	(2.510e-04)	(2.597e-04)	(2.686e-04)	(2.547e-04)	(2.405e-04)	(2.691e-04)
Age squared	−2.288e-05***	−2.252e-05***	−1.709e-05***	−2.060e-05***	−2.702e-05***	−1.975e-05***
	(1.333e-06)	(1.365e-06)	(1.346e-06)	(1.306e-06)	(1.228e-06)	(1.264e-06)
Forward number/age	4.088e-01***	4.079e-01***	4.132e-01***	4.057e-01***	4.010e-01***	4.034e-01***
	(4.167e-02)	(4.145e-02)	(4.059e-02)	(4.125e-02)	(4.340e-02)	(4.234e-02)
Number of inventors	2.326e-02**	2.388e-02**	2.051e-02*	2.248e-02**	2.268e-02**	1.826e-02*
	(1.327e-02)	(1.373e-02)	(1.335e-02)	(1.356e-02)	(1.184e-02)	(1.171e-02)
Number of IPCs	4.608e-02**	3.367e-02**	1.480e-01***	1.234e-02	1.242e-01***	3.464e-01***
	(2.040e-02)	(2.062e-02)	(2.144e-02)	(2.099e-02)	(2.045e-02)	(2.234e-02)
Number of applicants	−8.715e-02***	−9.311e-02***	−8.179e-02***	−8.601e-02***	−7.301e-02***	−6.248e-02***
	(7.764e-03)	(8.025e-03)	(8.051e-03)	(7.994e-03)	(7.315e-03)	(7.697e-03)
Number of claims	5.516e-03***	5.636e-03***	5.275e-03***	5.391e-03***	5.226e-03***	4.808e-03***
	(4.892e-04)	(4.853e-04)	(5.077e-04)	(5.112e-04)	(4.624e-04)	(4.743e-04)
Number of references	3.478e-03***	3.392e-03***	3.235e-03***	3.367e-03***	3.621e-03***	3.248e-03***
	(4.206e-04)	(4.128e-04)	(4.205e-04)	(4.151e-04)	(4.234e-04)	(4.293e-04)
References age mean	−2.018e-03***	−2.082e-03***	−1.604e-03***	−1.771e-03***	−2.098e-03***	−1.582e-03***
	(2.657e-04)	(2.671e-04)	(2.507e-04)	(2.491e-04)	(2.633e-04)	(2.441e-04)
References age variance	2.771e-06***	2.777e-06***	2.516e-06***	2.768e-06***	2.551e-06***	2.044e-06***
	(6.593e-07)	(6.620e-07)	(6.345e-07)	(6.406e-07)	(6.789e-07)	(6.511e-07)
Industry dummy (within IT)	2.012e-01***	1.794e-01***	2.976e-01***	1.004e-01***	3.494e-01***	6.093e-01***
	(1.631e-02)	(1.655e-02)	(1.573e-02)	(1.984e-02)	(1.981e-02)	(2.263e-02)
Industry dummy (across healthcare and IT)	2.727e-01***	2.625e-01***	1.315e-01***	2.306e-01***	7.568e-01***	9.211e-01***
	(1.658e-02)	(1.672e-02)	(1.787e-02)	(1.747e-02)	(3.348e-02)	(3.156e-02)
Log-likelihood	−10483223	−10479345	−10473816	−10480930	−10476333	−10444037
BIC	20966608	20958867	20947808	20962036	20952842	20888305

Notes. Values in parentheses are robust standard errors clustered by cited patents.
*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.
Industry dummy: within IT (1), across healthcare and IT (2), and within healthcare (reference group).

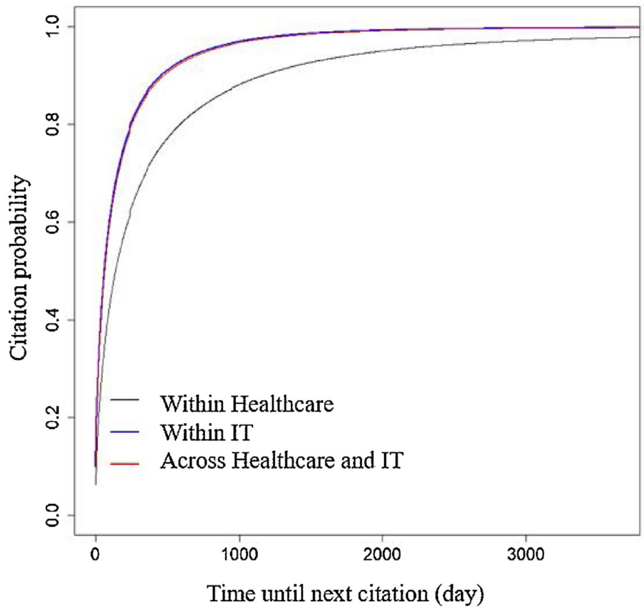

Fig. 1. Probability that next citation occurs within *t* days of elapsed time.

the citation rate. One possible explanation of this result is that technologies tend to be quickly replaced by subsequent technologies when the field is rapidly progressing (Haupt et al., 2007). In Model 3, we add *potential of technological convergence,* and this variable is also positive and significant. This intuitive result shows that a patent associated with a technology field that has a high potential of technology convergence is more likely to gain attention from a large number of actors. Third, *popularity of the technology field* is shown to be positive and significant in Model 4 as expected, whereas the result is mixed in the full model. That is, with the existence of other time-varying covariates, the relative influence of the popularity of the technology field seems to weaken. Lastly, Model 5 adds *technological novelty*, and this variable is shown to be positive and significant. Because the lower value of the minimum z-score represents higher novelty, this result means that extremely high novelty generally has a negative effect on the citation rate. This result can be partially explained by Kim et al. (2016), who showed that the majority of patents representing high novelty are likely to have a low impact, except for a small proportion that show high conventionality at the same time. By way of a robustness check, we compute the technological novelty by following the score constructed by Wang, Veugelers, and Stephan (2017) (see the Appendix A). Models 7 and 8 in the Appendix A show the robust result for technological novelty.

We can obtain a more valuable interpretation through scenario analysis, which can represent the patterns of a citation by the age of a patent under changing environmental conditions. Before showing the scenario analysis, we compare the overall results of the three groups analyzed. Fig. 1 shows the probability that the next citation of an arbitrary patent in each group occurs within days represented on the x axis, under the condition that all the other variables in the model are set at their mean value. We can see that the “within IT” and “across IT and healthcare” groups show similar probability patterns, while the “within healthcare” group has distinctively low probability values. This result shows that the citation pattern of convergence technologies created as a result of integrating heterogeneous knowledge bases (i.e., across IT and healthcare in our case) is not simply around the midpoint of the integrated areas but can sometimes be inclined toward one side, although this is not always the case.

Based on the results of the fitted Cox model, we conduct scenario analyses to understand how differences in the evolution of a technology field influence the citation patterns of a patent. For each environmental variable, we set three scenarios: its value is increasing, decreasing, or consistent as the focal patent ages. In each scenario that assumes an increasing influence, the field-level covariate is assumed to steadily increase from its mean to its maximum value. By contrast, the field-level covariate is assumed to gradually decrease from its mean to its minimum value in each scenario for a decreasing influence. Lastly, the remaining scenario assumes that the field-level covariate consistently retains its mean value. All the other variables are assumed to be their average values. Under such conditions, we compute the probability that the next citation of a focal patent will occur within one year, which is a tentative period set for our experiment, at each age of the focal patent.

Figs. 2–5 show the results of the scenario analyses for the four field-level dimensions considered. The subplots in each figure correspond to each class of industry dummy: (a) within healthcare (black), (b) within IT (blue), and (c) across healthcare and IT (red). The x axis of each plot indicates the age of the patent and the y axis indicates the probability that the next citation occurs within a year at the age indicated by the x axis. In addition, the three curves in each subplot correspond to the results obtained from the increasing, decreasing, and consistent scenarios.

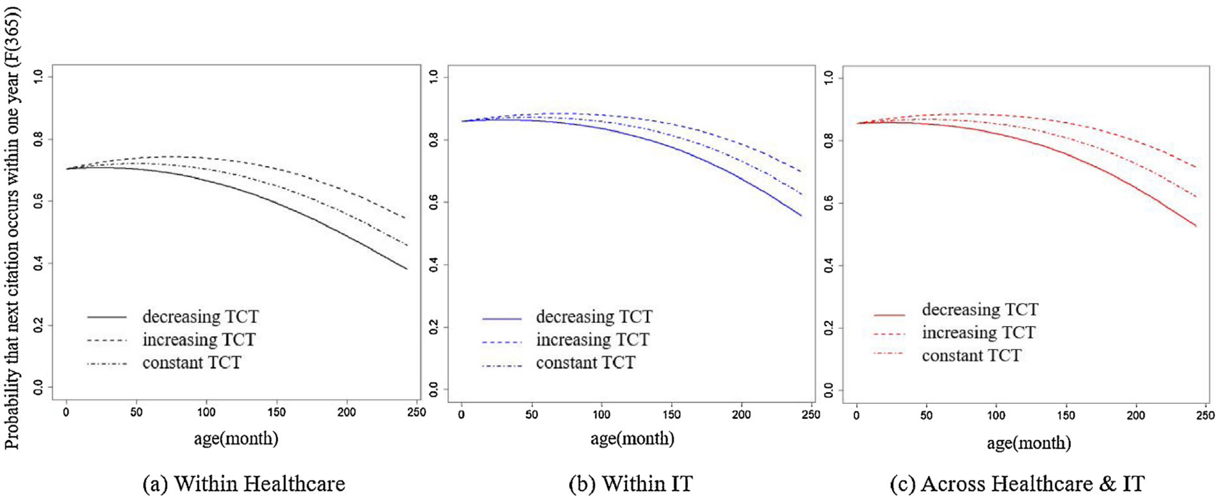

Fig. 2. Scenario analyses on the rate of technological progress.

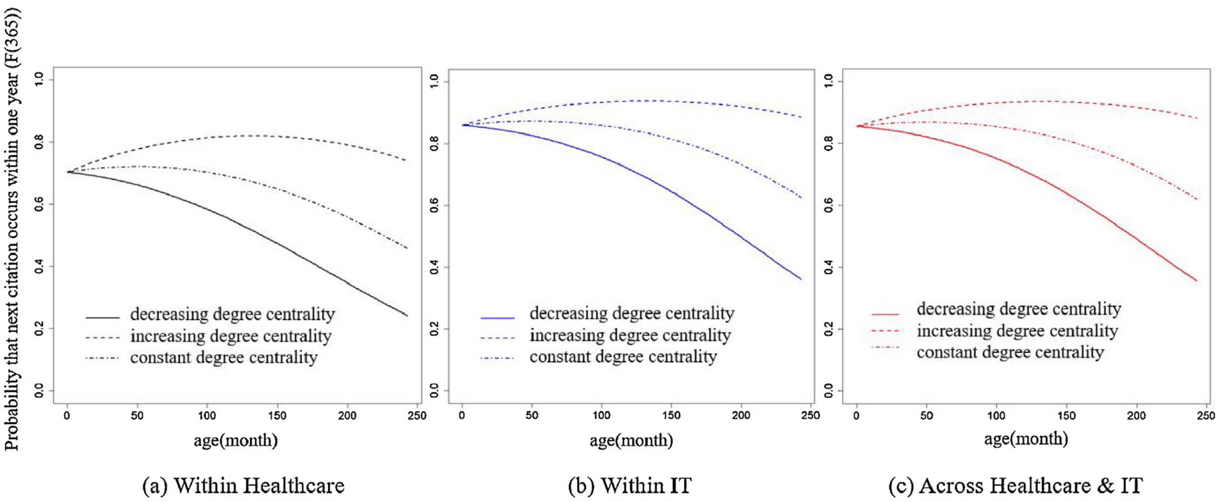

Fig. 3. Scenario analyses on the potential of technological convergence.

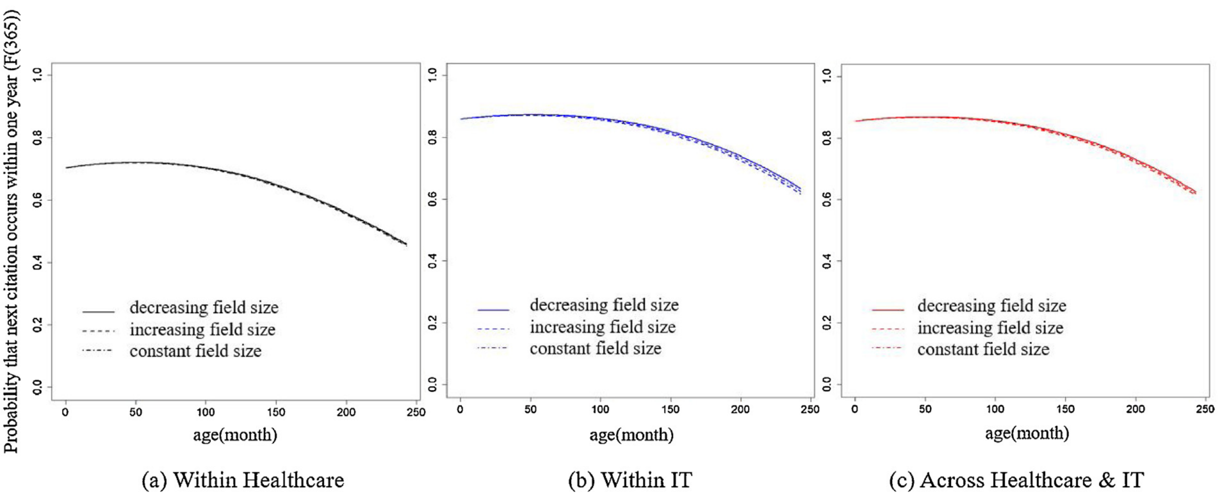

Fig. 4. Scenario analysis on the popularity of the technology field.

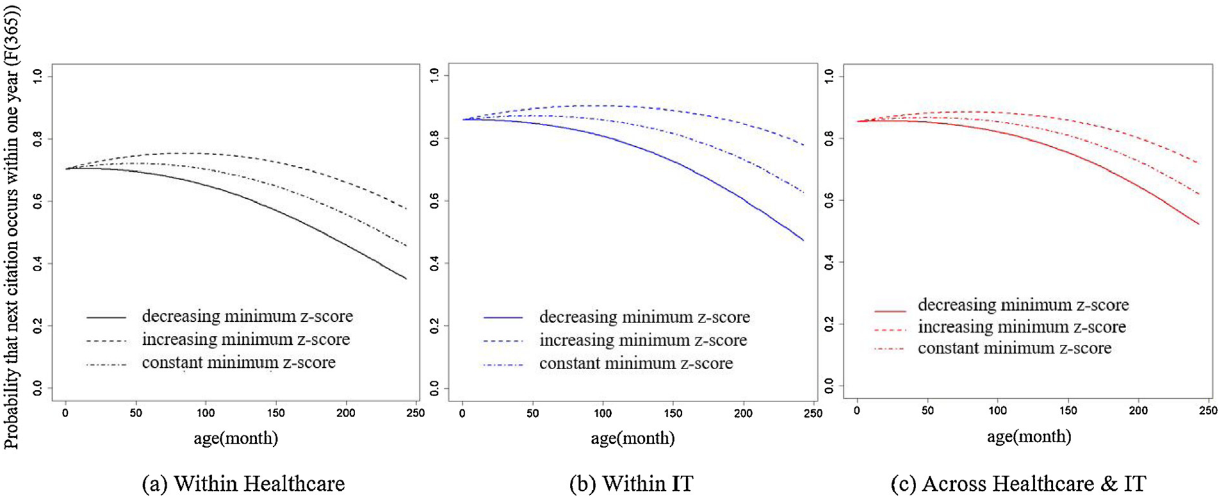

Fig. 5. Scenario analyses on the technological novelty.

First, Fig. 2 shows the results of the scenario analysis for TCT. Overall, there is no notable difference among the results in the three subplots ((a), (b), and (c)). In the decreasing TCT scenario, the rate of technological progress increases gradually. In the scenario in which the rate of technological progress increases (i.e., the decreasing TCT scenario), we find that the peak probability that the next citation occurs within a year appears at an earlier patent age than the scenario assuming a consistent or decreasing rate of technological progress. This finding shows that the impact of patents in technology fields in which the rate of progress increases tends to erode earlier than the impact of the other patents. In addition, the results show that the overall probability is lower in the decreasing TCT scenario than in the other scenarios. This result is partially attributable to the fact that cutting-edge technologies are quickly replaced when a technology field evolves from its introductory to its growth stage (Haupt et al., 2007). These results imply that unique patent citation patterns should be expected in each stage of the technology life cycle, as each stage displays different rates of technological progress. If a patent is applied for in the introductory stage of a technology field's evolution and the field is expected to gradually move toward a growth stage as the age of the focal patent increases, we can postulate that the rate of technological progress increases along with the age of the patent (i.e., the decreasing TCT scenario). When such a scenario is anticipated, the patent age when adoption most vigorously occurs (most citations) is expected to be lower than our general expectation. In addition, the overall probability of citation per patent might be lower than expected. Likewise, if a technology field is at the mid-point of its growth stage when a focal patent is introduced, the rate of technological progress is expected to decrease as a focal patent ages (i.e., the increasing TCT scenario). In such a scenario, the peak probability of receiving a citation may appear later than the expected age and the overall probability of citation per patent will be higher.

Second, Fig. 3 shows the results of the analysis focusing on the convergence potential of individual areas. This shows that the citation pattern of a patent is greatly influenced by different scenarios for the potential of technological convergence. If the convergence potential of the field increases gradually, not only is the peak age of a focal patent citation delayed, but the overall citation probability is also significantly enhanced. This result connotes the importance of decision makers' ability to sense the technology fields that will generate large amounts of convergence technologies in the future to avoid them underestimating the value of important patents. Indeed, recent research on network analysis has attempted to determine the relevant information using a method called link prediction (Kim et al., 2019; Park & Yoon, 2018). On the contrary, patents in fields entering the stage of diminishing convergence may lose their value much earlier than expected. Therefore, another crucial capability is detecting technology fields that have diminishing potential for technological convergence to avoid overestimating the value of the patents in such areas. Link prediction methods can also be applied to anticipate convergence links that are likely to disappear.

Third, Fig. 4 shows the results of the scenario analysis in terms of the popularity of the technology field. Interestingly, at least in our case analysis, changes in the popularity of the technology field do not result in a significant difference in the citation patterns of an individual patent. Although our analysis does not show whether the popularity of the technology field has a significant impact on citation patterns, one should take notice that the result may be influenced by the way we define the scope of the relevant technologies.

Lastly, Fig. 5 shows how the expected citation pattern of an individual patent changes with the decreasing or increasing technological novelty of the fields to which the focal patent belongs. The citation patterns of within-industry patents are affected by changes in technological novelty, whereas those lying across industries are relatively weakly affected by such changes. Moreover, when technological novelty erodes gradually (i.e., increasing the minimum z-score), the overall probability that the next citation occurs within a year rises and the peak point of this probability tends to appear later. The scenario in which technological novelty erodes gradually can be assumed when combined knowledge was unfamiliar at its early stage because of its outstanding novelty, but it will become more universal over time. The opposite scenario can be assumed when combined knowledge was popular at the time of the patent application, but this popularity has decreased because of unknown factors such as substitution by other technologies that have an identical function but better performance or the decline of customer demand for that technical function.

5. Conclusion

Despite a long history of examining technological change from the evolutionary perspective, studies using patent citation information have usually ignored how evolution in the technology field dynamically influences citation patterns. To investigate patent citations, most prior studies have relied on the innate features of the focal technology, thus overlooking the dynamic changes in the surrounding environment. This omission is surprising, as innovation scholars have recognized the unpredictability of technological change owing to complex and changing external conditions. In this study, we therefore suggested a model and scenario analysis to bridge this gap in the literature.

This study contributes to the literature and practice in three main ways. First, we provide a useful framework for decision makers who seek to understand the dynamics underlying patent citations. Our approach can be flexibly adjusted depending on the technological interests of the researcher or manager by changing the focal industries or supplementing key environmental aspects that evolve over time. Second, the scenario analyses based on the proposed model illustrated how different patent citation dynamics can be expected depending on the evolutionary stage and direction of the technology fields associated with the focal patent. Based on the expected scenario of macro-level technological evolution, one can evaluate a patent at hand in a more reliable way. For instance, if there is a considerable level of expected change in a particular environmental dimension, which seems to have relevance to the patent citation, one can experiment with the general influence of the particular dimension on patterns of patent citations. Through such an experiment, one can gain insights into the expected direction of changes in patterns of patent citations. This way of gaining foresight aligns with the evolutionary view, which ultimately argues that we can prepare for the future by understanding our history rather than making a precise prediction (Martin, 2010). Lastly, our framework will help decision makers not only in firms but also at public institutions by helping them gain a deeper understanding of the citation patterns of convergence technologies. Under the ongoing explosion of technological convergence, it has become increasingly difficult, or perhaps impossible in many areas, to anticipate the future impact of convergence technologies. Determining the conditions that shape the impact of convergence technologies has arisen as a key capability of organizations that govern technological investment decision making.

Along with these advantages, this study also has some limitations. First, it focused primarily on the environmental conditions related to technology. Future studies could reflect some key evolving environmental aspects other than technology such as organizational and market conditions. Second, our analysis essentially drew on past patent records. Although our suggestion has benefits for informing judgments on the direction of research and investment in pre-existing domains of convergence, such an approach may be difficult to employ for highly original convergence domains. Therefore, in future studies, it would be fruitful to design an approach that could help understand the impact of original convergence technologies.

Author contributions

Su Jung Jee: Conceived and designed the analysis, Collected the data, Performed the analysis, Wrote the paper. Minji Kwon: Collected the data, Performed the analysis. Jung Moon Ha: Collected the data, Performed the analysis. So Young Sohn: Conceived and designed the analysis, Wrote the paper (Supervision), Funding acquisition. A distinction is made between five types of contributions: Conceived and designed the analysis; Collected the data; Contribution data or analysis tools; Performed the analysis; Wrote the paper.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (2016R1A2A1A05005270).

Appendix A. Robustness check on technological novelty

For the robustness check on technological novelty, we use the maximum novelty score suggested by Wang et al. (2017). This score represents the maximum novelty among the journal pairs within the references of each paper. Because the authors construct the score from the journal papers, we make some modifications to suit the patent context. First, we use IPC pairs instead of journal pairs. Second, we consider every IPC pair instead of only new IPC pairs because a single patent may have only a few IPC pairs in it.

Suppose we want to obtain the maximum novelty score for a patent in year T. The following matrix provides the IPC co-occurrence profile (e.g., M_1 & M_2). The i- and j-th elements represent the number of times that both IPCs, M_i and M_j, co-occurred in the same patent in the preceding five years (i.e., T-4–T).

	M_1	M_2	M_3	M_4	M_5	…
M_1	/	0	3	0	2	…
M_2	0	/	12	3	5	…
M_3	3	12	/	2	0	…
M_4	0	3	2	/	0	…
M_5	2	5	0	0	/	…
…	…	…	…	…	…	/

The novelty score proposed by Wang et al. (2017) reflects the difficulty of making a particular knowledge combination. To obtain such information, the authors first assess the ease of making each combination using a cosine similarity index. In our context, the ease of making an IPC pair of M_1 and M_2 is defined as the cosine similarity between their IPC co-occurrence profiles:

$$COS_{1,2} = \frac{M_1 \cdot M_2}{\left||M_1|\right|\left||M_2|\right|}$$

where M_1 and M_2 are the row (or column) vectors.

Then, for each patent, the maximum novelty score is suggested as follows:

$$Novelty_{maxscore} = \max_{M_i \& M_j\, pair} (1 - COS_{i,j})$$

where the $M_i \& M_j$ pair includes possible IPC pairs within a patent and $(1 - COS_{i,j})$ indicates the difficulty of making the particular IPC pair.

Table A1 shows the results of the Cox regression, which is robust to the change in the novelty measure. A higher value of the maximum novelty score implies higher novelty (in the case of the minimum z-score, a lower value implies higher novelty).

Table A1
Results of the Cox regression.

	Model 7	Model 8
TCT (day)		8.068e-04*** (2.719e-05)
Potential of technological convergence (degree centrality)		2.377e-02*** (8.157e-04)
Popularity of technology field (number of patents)		9.530e-06*** (2.762e-06)
Technological novelty (cosine similarity based measure)	−1.834e-02*** (2.007e-02)	−3.048e-01*** (3.153e-02)
Age (month)	7.257e-04*** (2.511e-04)	−1.809e-04 (2.842e-04)
Age squared	−2.288e-05*** (1.333e-06)	−1.540e-05*** (1.378e-06)
Forward number/age	4.090e-01*** (4.170e-02)	4.131e-01*** (4.028e-02)
Number of inventors	2.327e-02** (1.327e-02)	2.098e-02* (1.411e-02)
Number of IPCs	4.720e-02*** (2.045e-02)	1.468e-01*** (2.216e-02)
Number of applicants	−8.712e-02*** (7.765e-03)	−8.875e-02*** (8.496e-03)
Number of claims	5.519e-03*** (4.891e-04)	5.420e-03*** (5.102e-04)
Number of references	3.481e-03*** (4.203e-04)	3.139e-03*** (4.095e-04)
References time lag mean	−2.024e-03*** (2.656e-04)	−1.682e-03*** (2.466e-044)
References time lag variance	2.750e-06*** (6.623e-07)	2.193e-06*** (6.423e-07)
Industry dummy (within IT)	2.008e-01*** (1.634e-02)	2.408e-01*** (2.183e-02)
Industry dummy (across healthcare and IT)	2.707e-01*** (1.687e-02)	6.145e-02*** (1.997e-02)
Log-likelihood	−10483218	−10467303
BIC	20966613	20934825

Notes. Values in parentheses are robust standard errors clustered by cited patents.
*p < 0.1, **p < 0.05, ***p < 0.01.
Industry dummy: within IT (1), across healthcare and IT (2), and within healthcare (reference group).

References

Ahuja, G., & Morris Lampert, C. (2001). Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal*, *22*(6-7), 521–543.

Albert, M. B., Avery, D., Narin, F., & McAllister, P. (1991). Direct validation of citation counts as indicators of industrially important patents. *Research Policy*, *20*(3), 251–259.

Alcacer, 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.

Alnuaimi, T., & George, G. (2016). Appropriability and the retrieval of knowledge after spillovers. *Strategic Management Journal*, *37*(7), 1263–1279.

Andersen, B. (1999). The hunt for S-shaped growth paths in technological innovation: A patent study. *Journal of Evolutionary Economics*, *9*(4), 487–526.

Banerjee, P. M., & Cole, B. M. (2011). Globally radical technologies and locally radical technologies: The role of audiences in the construction of innovative impact in biotechnology. *IEEE Transactions on Engineering Management*, *58*(2), 262–274.

Belderbos, R., Carree, M., & Lokshin, B. (2004). Cooperative R&D and firm performance. *Research Policy*, *33*(10), 1477–1492.

Belderbos, R., Cassiman, B., Faems, D., Leten, B., & Van Looy, B. (2014). Co-ownership of intellectual property: Exploring the value-appropriation and value-creation implications of co-patenting with different partners. *Research Policy*, *43*(5), 841–852.

Branstetter, L. G. (2001). Are knowledge spillovers international or intranational in scope?: Microeconometric evidence from the US and Japan. *Journal of International Economics*, *53*(1), 53–79.

Carpenter, M. P., Narin, F., & Woolf, P. (1981). Citation rates to technologically important patents. *World Patent Information*, *3*(4), 160–163.

Chen, L. (2017). Do patent citations indicate knowledge linkage? The evidence from text similarities between patents and their citations. *Journal of Informetrics*, *11*(1), 63–79.

Curran, C. S., & Leker, J. (2011). Patent indicators for monitoring convergence: Examples from NFF and ICT. *Technological Forecasting and Social Change*, *78*(2), 256–273.

Davies, A. (1996). Innovation in large technical systems: The case of telecommunications. *Industrial and Corporate Change*, *5*(4), 1143–1180.

Dosi, G. (1982). Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Research Policy*, *11*(3), 147–162.

Duysters, G., & Hagedoorn, J. (1998). Technological convergence in the IT industry: The role of strategic technology alliances and technological competencies. *International Journal of the Economics of Business*, *5*(3), 355–368.

Fischer, T., & Leidinger, J. (2014). Testing patent value indicators on directly observed patent value: An empirical analysis of Ocean Tomo patent auctions. *Research Policy*, *43*(3), 519–529.

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science*, *47*(1), 117–132.

Fung, M. K., & Chow, W. W. (2002). Measuring the intensity of knowledge flow with patent statistics. *Economics Letters*, *74*(3), 353–358.

Geum, Y., Kim, C., Lee, S., & Kim, M. S. (2012). Technological convergence of IT and BT: Evidence from patent analysis. *ETRI Journal*, *34*(3), 439–449.

Gilfillan, S. C. (1935). *Inventing the ship*. Chicago: Follett Publishing Company.

Hacklin, F. (2007). *Management of convergence in innovation: Strategies and capabilities for value creation beyond blurring industry boundaries*. Heidelberg: Physica.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*, *36*(1), 16–38.

Harhoff, D., Narin, F., Scherer, F. M., & Vopel, K. (1999). Citation frequency and the value of patented inventions. *The Review of Economics and Statistics*, *81*(3), 511–515.

Haupt, R., Kloyer, M., & Lange, M. (2007). Patent indicators for the technology life cycle development. *Research Policy*, *36*(3), 387–398.

Hausman, J. A., Hall, B. H., & Griliches, Z. (1984). *Econometric models for count data with an application to the patents-R&D relationship*. National Bureau of Economic Research.

He, Z. L., & Deng, M. (2007). The evidence of systematic noise in non-patent references: A study of New Zealand companies' patents. *Scientometrics*, *72*(1), 149–166.

Hu, M. C., & Mathews, J. A. (2008). China's national innovative capacity. *Research Policy*, *37*(9), 1465–1479.

Huang, Y. F., & Chen, C. J. (2010). The impact of technological diversity and organizational slack on innovation. *Technovation*, *30*(7-8), 420–428.

Jaffe, A. B., & De Rassenfosse, G. (2017). Patent citation data in social science research: Overview and best practices. *Journal of the Association for Information Science and Technology*, *68*(6), 1360–1374.

Jaffe, A. B., & Trajtenberg, M. (1996). Flows of knowledge from universities and federal laboratories: Modeling the flow of patent citations over time and across institutional and geographic boundaries. *Proceedings of the National Academy of Sciences*, *93*(23), 12671–12677.

Jaffe, A. B., Trajtenberg, M., & Fogarty, M. S. (2000). Knowledge spillovers and patent citations: Evidence from a survey of inventors. *The American Economic Review*, *90*(2), 215–218.

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

Jee, S. J., & Sohn, S. Y. (2015). Patent network based conjoint analysis for wearable device. *Technological Forecasting and Social Change*, *101*, 338–346.

Karvonen, M., & Kässi, T. (2011). Patent analysis for analysing technological convergence. *Foresight*, *13*(5), 34–50.

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.

Kayal, A. A., & Waters, R. C. (1999). An empirical evaluation of the technology cycle time indicator as a measure of the pace of technological progress in superconductor technology. *IEEE Transactions on Engineering Management*, *46*(2), 127–131.

Kim, N., Lee, H., Kim, W., Lee, H., & Suh, J. H. (2015). Dynamic patterns of industry convergence: Evidence from a large amount of unstructured data. *Research Policy*, *44*(9), 1734–1748.

Kim, D., Cerigo, D. B., Jeong, H., & Youn, H. (2016). Technological novelty profile and invention's future impact. *EPJ Data Science*, *5*(1), 8.

Kim, J., Kim, S., & Lee, C. (2019). Anticipating technological convergence: Link prediction using Wikipedia hyperlinks. *Technovation*, *79*, 25–34.

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science*, *3*(3), 383–397.

Kuhn, T. S. (1963). . *The structure of scientific revolutions* (Vol. 2) Chicago: University of Chicago Press.

Lanjouw, J. O., & Schankerman, M. (1999). *The quality of ideas: Measuring innovation with multiple indicators (No. w7345)*. National Bureau of Economic Research.

Lavie, D., & Rosenkopf, L. (2006). Balancing exploration and exploitation in alliance formation. *The Academy of Management Journal*, *49*(4), 797–818.

Lee, Y. G., Lee, J. D., Song, Y. I., & Lee, S. J. (2007). An in-depth empirical analysis of patent citation counts using zero-inflated count data model: The case of KIST. *Scientometrics*, *70*(1), 27–39.

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

Marco, A. C. (2007). The dynamics of patent citations. *Economics Letters*, *94*(2), 290–296.

Martin, B. R. (2010). The origins of the concept of 'foresight' in science and technology: An insider's perspective. *Technological Forecasting and Social Change*, *77*(9), 1438–1447.

Mukherjee, S., Romero, D. M., Jones, B., & Uzzi, B. (2017). The nearly universal link between the age of past knowledge and tomorrow's breakthroughs in science and technology: The hotspot. *Science Advances*, *3*(4), e1601315.

Mukherji, N., & Silberman, J. (2013). Absorptive capacity, knowledge flows, and innovation in US metropolitan areas. *Journal of Regional Science*, *53*(3), 392–417.

Nalebuff, B. J., Brandenburger, A., & Maulana, A. (1996). *Co-opetition*. London: Harper Collins Business.

Narin, F. (1993). Technology indicators and corporate strategy. *Review of Business*, *14*(3), 19.
Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Cambridge: Harvard University Press.
Nelson, R. R., Dosi, G., Helfat, C. E., Pyka, A., Saviotti, P. P., Lee, K., et al. (2018). *Modern evolutionary economics: An overview*. Cambridge: Cambridge University Press.
Nerkar, A., & Paruchuri, S. (2005). Evolution of R&D capabilities: The role of knowledge networks within a firm. *Management Science*, *51*(5), 771–785.
Park, I., & Yoon, B. (2018). Technological opportunity discovery for technological convergence based on the prediction of technology knowledge flow in a citation network. *Journal of Informetrics*, *12*(4), 1199–1222.
Petruzzelli, A. M., Rotolo, D., & Albino, V. (2015). Determinants of patent citations in biotechnology: An analysis of patent influence across the industrial and organizational boundaries. *Technological Forecasting and Social Change*, *91*, 208–221.
Podolny, J. M., & Stuart, T. E. (1995). A role-based ecology of technological change. *The American Journal of Sociology*, *100*(5), 1224–1260.
Preschitschek, N., Niemann, H., Leker, J. G., & Moehrle, M. (2013). Anticipating industry convergence: Semantic analyses vs IPC co-classification analyses of patents. *Foresight*, *15*(6), 446–464.
Reagans, R., & Zuckerman, E. W. (2001). Networks, diversity, and productivity: The social capital of corporate R&D teams. *Organization Science*, *12*(4), 502–517.
Reitzig, M. (2003). What determines patent value? Insights from the semiconductor industry. *Research Policy*, *32*(1), 13–26.
Rosenberg, N. (1963). Technological change in the machine tool industry, 1840–1910. *The Journal of Economic History*, *23*(4), 414–443.
Simonton, D. K. (1999). *Origins of genius: Darwinian perspectives on creativity*. New York: Oxford University Press.
Stolpe, M. (2002). Determinants of knowledge diffusion as evidenced in patent data: The case of liquid crystal display technology. *Research Policy*, *31*(7), 1181–1198.
Strumsky, D., & Lobo, J. (2015). Identifying the sources of technological novelty in the process of invention. *Research Policy*, *44*(8), 1445–1461.
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.
Touati, F., & Tabish, R. (2013). U-healthcare system: State-of-the-art review and challenges. *Journal of Medical Systems*, *37*(3), 9949.
Tong, X., & Frame, J. D. (1994). Measuring national technological performance with patent claims data. *Research Policy*, *23*(2), 133–141.
Trajtenberg, M. (1990). A penny for your quotes: Patent citations and the value of innovations. *The Rand Journal of Economics*, *21*(1), 172–187.
Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, *342*(6157), 468–472.
Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The small world problem. *The American Journal of Sociology*, *111*(2), 447–504.
Van den Berghe, L., & Verweire, K. (2001). Convergence in the financial services industry. *The Geneva Papers on Risk and Insurance-Issues and Practice*, *26*(2), 173–183.
Verhoeven, D., Bakker, J., & Veugelers, R. (2016). Measuring technological novelty with patent-based indicators. *Research Policy*, *45*(3), 707–723.
von Delft, S. (2013). Inter-industry innovations in terms of electric mobility: Should firms take a look outside their industry? *Journal of Business Chemistry*, *10*(2), 67–87.
Wang, J., Veugelers, R., & Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, *46*(8), 1416–1436.
WIPO. (2019). *International patent classification: Frequently asked questions*. http://www.wipo.int/classifications/ipc/en/faq/

Su Jung Jee is a PhD candidate at the Department of Industrial Engineering, Yonsei University. Her research areas include innovation management and policy.

Minji Kwon participated in this research while she was a student at the Department of Industrial Engineering, Yonsei University.

Jung Moon Ha participated in this research while he was a student at the Department of Industrial Engineering, Yonsei University.

So Young Sohn is a Professor at the Department of Industrial Engineering at Yonsei University in Korea. Her research areas include technology management, marketing, quality, and reliability engineering. Detailed information about her teaching and research areas can be found at http://isl.yonsei.ac.kr.