Information Research - Vol. 31 No. iConf (2026)

A study on the correlation mechanism between knowledge convergence characteristics and short- and long-term patent impact

Ma Jing, Li Liyuan, and Pan Yaohui

DOI: https://doi.org/10.47989/ir31iConf64126

Abstract

Introduction. This study aims to reveal the underlying mechanisms linking knowledge convergence characteristics in different technological domains with the short- and long-term impact of patents.

Method. Using a negative binomial regression model, it systematically examines the differentiated effects of scientific knowledge breadth, scientific knowledge depth, technological knowledge breadth, and technological knowledge depth on the short- and long-term impact of patents.

Analysis. This framework enables precise characterization of structural features in knowledge absorption and recombination and provides the basis for quantifying heterogeneous effects of knowledge convergence.

Results. The breadth of knowledge generally enhances patent influence and exhibits an inverted U-shaped relationship, suggesting that moderate diversity and integration increase patent value. In contrast, the effect of knowledge depth varies across fields and time scales: scientific depth in computer technology shows a long-term U-shaped pattern, depth in semiconductors has a positive long-term effect, while technological depth in biotechnology is largely negative.

Conclusion. The innovation cycles and knowledge application mechanisms in different fields determine the configuration patterns of knowledge breadth and depth: fast-iteration fields rely more on short-term diffusion driven by breadth, while long-cycle fields emphasize the long-term value of deep accumulation.

Introduction

In the context of a globalized knowledge economy, technological competition and innovation capabilities have become the core driving force for industrial upgrading and economic growth. As a key measure of technological competitiveness and innovation capabilities, patent influence profoundly reflects the actual role and potential value of a patent in subsequent technological innovation, industry development, and market competition (Haessler et al., 2023). The technological value has the characteristics of cross-cycle and continuous evolution. This dynamic attribute determines that relying solely on a static perspective to evaluate patent influence has obvious limitations and needs to be optimized in the time dimension. Specifically, patent influence can be discussed from two levels: short-term and long-term. Short-term influence is mainly reflected in the immediate application value of patents in technology market competition, while long-term influence more concentratedly reflects its lasting contribution to technological progress and knowledge accumulation (Gan & Zhang, 2025). Knowledge convergence is regarded as an important driving force for patent innovation capabilities and influence. Its essence is the structural characteristics formed by the mutual combination and reorganization of different technical fields, knowledge elements, or research paths in the patent innovation process (Liu, Mao, & Guan, 2020).

Existing research shows that knowledge convergence can enrich the technical background of patents, so that patents not only have technical advantages in specific fields, but also absorb innovative achievements in other fields, form cross-domain knowledge integration, enhance the breadth and depth of technology, and thus provide more possibilities for technology diffusion and re-innovation (Li & Wang, 2024). However, the effects of knowledge convergence on the short-term and long-term impact of patents may differ: in the short term, the high degree of integration of cross-disciplinary knowledge may increase the complexity of knowledge absorption and application, thereby limiting the immediate market transformation and short-term technological impact of patents; in the long term, cross-disciplinary knowledge, by enhancing technological scalability and knowledge spillover effects, is more likely to promote continuous technological accumulation and innovation diffusion, significantly improving long-term technological influence (Xu et al., 2025).

While numerous studies have focused on the relationship between knowledge convergence and patent influence, few systematically examine its heterogeneous effects from both short-term and long-term perspectives. In particular, patent lifecycle structures and knowledge update rhythms vary significantly across different technological fields, leading to the possibility that the same convergence pattern can have opposite effects at different stages. Therefore, this study focuses on further exploring the specific mechanisms by which different types of knowledge convergence characteristics, including the breadth and depth of scientific and technological knowledge absorption, influence patents' short-term and long-term impacts. Furthermore, this study examines the differences in these effects across different technological fields, thereby revealing the underlying mechanisms by which cross-domain knowledge convergence influences patent value.

This study focuses on knowledge-intensive technological fields characterized by frequent patent activity and rapid knowledge iteration, exploring the underlying mechanisms between knowledge convergence characteristics and patents' short-term and long-term influence. By constructing a structured knowledge convergence measurement metric, this study aims to reveal the patterns and mechanisms of patent knowledge convergence across different fields. Patent networks serve as the foundation for a structured analysis of patent knowledge convergence and patent influence. This study uses patents as core nodes and utilizes scientific papers and patent literature to characterize the sources of knowledge convergence within scientific and technological fields. At the network structural level, the direction of citation relationships is considered to characterize the knowledge flow paths between patents. At the node level, based on network construction, this study introduces knowledge similarity weights to improve the accuracy of measuring the knowledge distance between different technical topics during the knowledge convergence process, thereby characterizing the inherent characteristics of patents in the process of knowledge absorption and recombination. To further reveal the inherent relationship between knowledge convergence characteristics and patent influence, this study constructs a quantitative model for short-term and long-term patent influence based on the consideration of patent characteristics and knowledge convergence characteristics, systematically quantifying the impact path of knowledge convergence characteristics on patent influence. In view of the discrete distribution characteristics of influence data, this study uses a negative binomial regression model for quantitative analysis to identify the significance and marginal effects of different types of knowledge convergence on patents in the short and long term.

Related work

Patent citation and knowledge relevance

Patent citation is one of the key indicators for characterizing the flow of technological knowledge (Wang et al., 2013), and its research paradigms and application value continue to deepen and expand in the academic field. At the macro level, existing research focuses on the application of patent citations in tracing the path of technological evolution (Bhatt et al., 2023; Oh et al., 2023) and building patent value assessment systems (Hu, Zhou, & Lin, 2023; Hong, Kim, Woo, Kim, & Lee, 2022), revealing the knowledge diffusion mechanism based on documents through citation networks (Fernández et al., 2022; Pan et al., 2024).

Research on tracing knowledge flows through citation has gone through several stages of development. Early researchers mainly relied on numerical indicators such as the number of citations to characterize knowledge flows and assumed that all citations had the same effect. This method, which relies solely on statistical data, has obvious limitations, and fails to fully consider the complexity of citation behavior (Wang et al., 2020). With the development of natural language processing and text mining technology, scholars began to pay attention to the knowledge attributes of citation relationships. Chen (2017) used the WF-IDF weighted vector space model (VSM) to systematically verify the technological relevance of citation behavior by quantifying the similarity of patent texts. However, since authors usually rewrite the content when citing documents, it is difficult to accurately measure the semantic relevance between cited documents and citing documents. To this end, researchers have gradually adopted analysis methods based on text semantics. In the early days, Word2Vec and Glove were the representatives, while at present, BERT and its derivative models have become mainstream tools due to their stronger context understanding ability (Hou et al., 2023).

In order to reveal the internal mechanism of citation propagation, the prior knowledge is further classified. Studies have shown that technical knowledge from patents forms a gradual technical iteration through the patent citation network, achieving optimization and improvement of existing technologies. At the same time, scientific knowledge from papers also plays an important role in promoting patent innovation. For example, Ke and Qing (2020) found that basic scientific papers and novel papers are more likely to have a direct technical impact on patents. In addition, Kong et al. (2023) combined the dual sources of knowledge of science and technology and explored the influence mechanism on patents by measuring factors such as the intensity and breadth of knowledge absorbed by cited patents.

In recent years, research has further extended to micro-mechanisms, starting from the internal motivation and knowledge structure of patent citation behavior, and exploring the differentiated role of different types of knowledge in the citation process. Hou et al. (2023) revealed the differentiated influence of different categories of prior knowledge such as theoretical derivation and experimental methods in the process of technology diffusion through semantic similarity analysis. These studies indicate that patent citation research is shifting from traditional network structure analysis to knowledge content assessment.

Knowledge convergence measurement of patented technologies

In today's rapidly developing science and technology, innovation activities are increasingly dependent on the cross-integration of different knowledge fields and technology systems. As the boundaries between disciplines and industries gradually blur, major technological breakthroughs are often no longer limited to a single field but emerge at the intersection of multiple technologies. As an important carrier of technological knowledge, the knowledge convergence characteristics of patents have gradually become the focus of academic attention. Its research mainly focuses on the representation of knowledge diversity in citation networks and its structural effects. At the theoretical framework level, knowledge convergence research is usually based on the association pattern between knowledge units, revealing the integration mechanism and innovation potential between technologies through co-occurrence, coupling, and co-evolution. Nakamura et al. (2015) developed the DB-combination model, systematically distinguishing three types of knowledge combination patterns: depth (D), transfer (T), and breadth (C), and verified the feasibility of cross-industry technology association through patent data. Focusing on the patent field, the method of measuring knowledge convergence presents a dual path of network analysis and semantic mining. Huang et al. (2022) pointed out that patent citation relationships construct the inheritance and diffusion path of technical knowledge through forward citation and backward citation. Battke et al. (2016) proposed that patents with high knowledge diversity are more likely to occupy key hub positions in the citation network by integrating heterogeneous technical elements. Their intermediary centrality advantage as '*gatekeepers*' is positively correlated with the ability of cross-border knowledge diffusion.

In terms of methodological innovation, Choi & Yoon (2022) used network embedding methods and citation analysis to quantify the knowledge convergence distance of a single patent and used this knowledge convergence distance to quantify the degree of technical association between patents. Gao & Jiang (2024). characterized the convergence of technical knowledge from the semantic perspective of patent texts. By combining topic models with document embedding models, they identified convergence topics across technical fields at the semantic level and further quantified their diversity, homogeneity, and cohesion. Lai and Su (2024) propose a comprehensive framework that bridges the knowledge flow perspective with the roles of knowledge sources and recipients in shaping technology convergence. Their study systematically classifies technology convergence patterns by examining how knowledge attributes and inter-firm collaborations influence cross-disciplinary convergence processes, highlighting how mature and dynamic knowledge flows can accelerate or diversify technological integration across fields.

In summary, existing research has achieved rich results in the areas of citation and knowledge association, and knowledge convergence measurement of patent technology. It has gradually moved from macro-network structure analysis to the characterization of knowledge attributes of relationships and has achieved a methodological expansion from traditional statistical methods to text mining and semantic modeling. However, there are still two deficiencies in current research: first, the discussion of knowledge convergence and patent influence remains more at the overall level, lacking a systematic separation and comparison of short-term and long-term effects; second, the heterogeneous effects of different types of knowledge convergence characteristics in various technical fields have not been fully revealed. Based on this, this article will be based on the field of knowledge-intensive technology, focusing on the breadth and depth characteristics of scientific knowledge and technological knowledge, to construct patent knowledge convergence measurement indicators, and combine patent networks to explore the mechanism and differences of its influence on patents from the two-time dimensions of short-term and long-term.

Data and research methods

Data sources and preprocessing

The patent data in this study is derived from the PatentsView data platform, an open database supported by the United States Patent and Trademark Office that provides basic information on patents and citation relationship data between patents. Regarding the citation data of patents to scientific papers, this study applies the ROSD dataset (Marx & Fuegi, 2020; 2022) and OpenAlex database is used to obtain detailed information on the cited papers. The specific data collection process is shown in Figure 1.

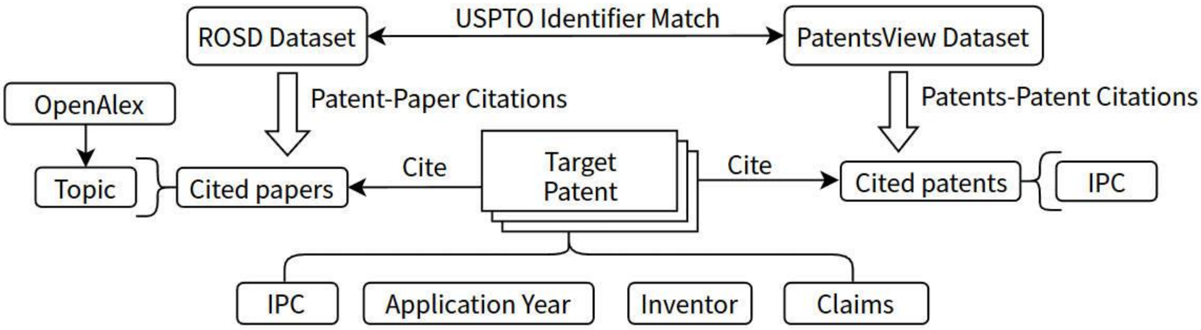

Figure 1. Data collection process.

To analyze knowledge absorption patterns across different technological fields, biotechnology, semiconductors, and computer technology—three representative fields characterized by high patent activity and rapid knowledge evolution are selected. Given that the International Patent Classification (IPC) has been established since 1971 and that this study primarily examines citations within 20 years of patent application, the patent application window has been ultimately set to 1971 to 2001. Finally, relevant patents in the three fields are screened using the WIPO technology classification standards (Szczygielski & Mycielski, 2024), retaining only those that cited both scientific papers and other patents. After data cleaning and matching, a total of 114,381 valid data samples have been retrieved. The sample distribution by field is shown in Table 1.

WIPO Code	Technological Field	Number of Patents
6	Computer Technology	53,200
8	Semiconductors	26,588
15	Biotechnology	34,593

Table 1. Distribution of patents by technological field, 1971–2001.

Construction of relevant indicators

Knowledge convergence

This study measures the diversity of knowledge absorbed by patents through the breadth of scientific knowledge and the breadth of technological knowledge. Existing calculation methods usually directly count the number of patent-cited papers and the fields of cited patents, and do not give sufficient consideration to the correlation between technologies between fields or subject classifications. Based on this, this study introduces a new similarity devaluation mechanism to the construction of knowledge breadth. The core idea is that if the newly cited paper topic or patent IPC is similar to the counted paper topic or patent IPC, its new diversity contribution is reduced. When a paper topic or patent IPC is cited multiple times, its contribution is repeatedly calculated based on similarity, but the new contribution decreases each time. The specific calculation method is

$$S\_Breadth=\sum_{i=1}^{N} (1-\max_{j<i} w(t_i,t_j)\times\frac{c_j}{N}) \quad (1)$$

Where N represents the total number of paper citations by the patent, $w(t_i,t_j)$ is the similarity weight between the i-th and j-th cited paper topics, and c_j is the cumulative number of citations belonging to topic j,

$$T\_Breadth=\sum_{i=1}^{N} (1-\max_{j<i} w(t_i,t_j)\times\frac{c_j}{N}) \quad (2)$$

Where N represents the total number of patent citations, $w(t_i,t_j)$ is the similarity weight between the i-th and j-th cited IPCs, andc_j is the cumulative citation count of IPC j.

In addition to measuring the diversity of knowledge, this study also uses two indicators, scientific knowledge depth and technical knowledge depth, to measure the concentration of scientific and technical knowledge absorbed by patents. Based on the traditional Herfindahl Index (HHI) (Moorthy & Polley, 2010), this study combines the number of citations of similar topics or similar IPCs to calculate the effective concentration. The specific calculation method is

$$S\_Depth=\sum_{k=1}^{K} (\frac{\sum_{t\in T_k} n_t}{N})^2 \quad (3)$$

where $T_k$denotes the k-th aggregated category (topics merged when similarity ≥ 0.8), n_t is the citation count of topic t, and N is the total number of paper citations.

$$T\_Depth=\sum_{k=1}^{K} (\frac{\sum_{t\in T_k} n_t}{N})^2 \quad (4)$$

where T_k denotes the k -th aggregated category (IPCs merged when similarity ≥ 0.8), n_t is the citation count of IPC t, and N is the total number of patent citations.

In addition, for the calculation of similarity weight, the study used co-occurrence analysis to quantify the similarity between topics or IPCs, and the specific calculation method is:

$$w(t_i,t_j)=\frac{c}{\sqrt{a*b}} \quad (5)$$

where c is the number of co-occurrences between topics or IPCs i and j, a is the occurrence count of topic or IPC i, and b is that of j.

Patent impact

Patent impact is primarily measured by the number of forward citations (Zhang, Qian, Huang, Guo, Zhang, & Lu, 2017), considering both short-term and long-term impact. Short-term impact is measured by forward citations within five years after application, reflecting the patent's early academic and technological influence,

$$S\_Impact=\sum_{t}^{t+4} C_i,t \quad (6)$$

where t is the application year of patent i, and C_i,t is the number of forward citations received in year t.

Long-term impact is measured by forward citations within 5 to 20 years after application, reflecting sustained influence and long-term technological value,

$$L\_Impact=\sum_{t+5}^{t+19} C_i,t \quad (7)$$

where t is the application year of patent i, and C_i,t is the number of forward citations received in year t.

This dual-perspective evaluation captures both the timeliness and persistence of patent influence, avoiding the limitations of single time-dimension measurement and providing a more comprehensive assessment of patent value.

Other variables

This study focuses on the core logic and influencing mechanisms of patent knowledge absorption, combining research paradigms in the field of technological innovation to construct a variable system from two perspectives: knowledge absorption and time effects. Furthermore, this article introduces several control variables to better control for potential influencing factors. First, the '*Year*' is used to control for the impact of the technological development stage, accounting for differences in technological paradigms and innovation activity across different eras. Second, considering the independent role of teamwork on patent influence, the '*Inventors*' is introduced to reflect the size of the R&D team. Finally, the '*Claims*' is used as an indicator to measure patent quality and scope of protection, which is generally positively correlated with patent drafting quality and technological coverage (Li et al., 2024). A full list of variables involved in this study is shown in Table 2.

No.	Type	Variable	Description	Calculation
1	Independent	S_Breadth	Diversity of absorbed scientific knowledge	Formula (1)
2		T_Breadth	Diversity of absorbed technological knowledge	Formula (2)
3		S_Depth	Concentration of absorbed scientific knowledge	Formula (3)
4		T_Depth	Concentration of absorbed technological knowledge	Formula (4)
5	Dependent	S_Impact	Early academic and technological influence	Formula (6)
6		L_Impact	Sustained impact and long-term value	Formula (7)
7	Control	Year	Application year of the patent	Patent application year
8		Inventors	Total number of inventors	Number of inventors in patent data
9		Claims	Number of claims, proxy for quality and protection scope	Number of claims in patent data

Table 2. Description of variables.

Subsequent analysis will conduct field-specific validation in biotechnology, semiconductors, and computer technology, in order to examine the heterogeneity of mechanisms across domains and refine the analysis based on the knowledge flow characteristics of each field.

Experiments and results analysis

Descriptive statistics and correlation analysis

To understand the basic distribution characteristics of the data, this study first conducts a descriptive statistical analysis of each variable. Specific statistical indicators include mean, standard error, minimum value, and maximum value. As can be seen from Table 3, the three technical fields show a certain degree of difference in the concentration and diversity of scientific and technological knowledge absorbed. For example, in terms of the breadth of scientific knowledge absorption, the mean value of biotechnology is significantly higher than that of computer technology and semiconductors, indicating that biotechnology patents cover more diverse topics when citing literature and have stronger knowledge integration capabilities. In terms of the depth of scientific knowledge absorption, the mean value of biotechnology is much lower than that of the other two fields, which means that semiconductors and computer technology have stronger '*deep cultivation*' characteristics in the process of scientific research knowledge absorption. In addition, in terms of the breadth and depth of technical knowledge, the three fields also show different citation strategy preferences, reflecting the heterogeneity of the knowledge accumulation and innovation models behind them.

A closer look at patent impact indicators reveals significant differences across fields in both early and long-term impact. The biotechnology has the lowest mean S_Impact, indicating that its patents have relatively limited early-stage academic and technological impact. However, its L_Impact has a wide dispersion, suggesting that some patents may possess strong scalability and breakthrough potential in terms of long-term value. In contrast, the computer technology exhibits relatively high mean values for both S_Impact and L_Impact, demonstrating that its patents not only have a rapid impact in the early stages but also possess strong, sustained impact over the long term. The semiconductors fall somewhere in between these two dimensions, with a mean S_Impact higher than that of biotechnology but lower than that of computer technology. Its mean L_Impact also demonstrates some signs of long-term value accumulation. Regarding control variables, the mean number of claims in the computer technology is higher than that of both biotechnology and semiconductors, indicating a greater emphasis on the boundaries of technological protection during patent drafting. The number of inventors varies little across fields, with a mean close to 1, indicating that R&D models primarily based on individuals or small teams predominate in the sample.

Field	Variable	VIF	Mean	Std. Error	Min	Max
Computer Technology	S_Breadth	1.482	10.253	12.483	1.000	61.704
	T_Breadth	1.103	9.595	6.787	0.000	22.765
	S_Depth	1.419	0.256	0.156	0.014	0.696
	T_Depth	1.029	0.686	0.280	0.000	1.000
	S_Impact	—	11.909	17.205	0.000	486.000
	L_Impact	—	37.236	71.822	0.000	2320.000
	Inventors	1.001	1.022	0.161	1.000	6.000
	Claims	1.051	21.061	17.037	1.000	393.000
Semiconductors	S_Breadth	1.519	8.061	9.612	1.000	61.704
	T_Breadth	1.162	8.128	6.417	0.000	22.765
	S_Depth	1.422	0.262	0.125	0.027	0.696
	T_Depth	1.048	0.706	0.284	0.000	1.000
	S_Impact	—	8.232	12.062	0.000	239.000
	L_Impact	—	22.194	50.682	0.000	3751.000
	Inventors	1.001	1.027	0.188	1.000	5.000
	Claims	1.095	16.173	13.570	1.000	300.000
Biotechnology	S_Breadth	1.706	39.613	22.624	1.000	61.704
	T_Breadth	1.446	4.658	4.980	0.000	22.765
	S_Depth	1.652	0.097	0.097	0.008	0.696
	T_Depth	1.385	0.576	0.307	0.000	1.000
	S_Impact	—	3.515	6.547	0.000	191.000
	L_Impact	—	19.525	45.274	0.000	1998.000
	Inventors	1.006	1.079	0.311	1.000	6.000
	Claims	1.059	17.098	17.133	1.000	683.000

Table 3. VIF values and descriptive statistics.

Before conducting regression analysis, a variance inflation factor (VIF) test has been conducted on the variables in each technical field to assess whether there is multicollinearity between the variables. The test results in Table 3 show that the VIF values of the variables in each field are all less than 2, indicating a low risk of multicollinearity between the variables, meeting the prerequisites for subsequent regression analysis. However, given that multicollinearity may still exist when the VIF value is less than 10, the potential impact of this risk on the regression analysis cannot be completely ruled out. Therefore, to ensure the robustness and explanatory power of the model, the correlation coefficients is calculated between the variables and a heat map is created.

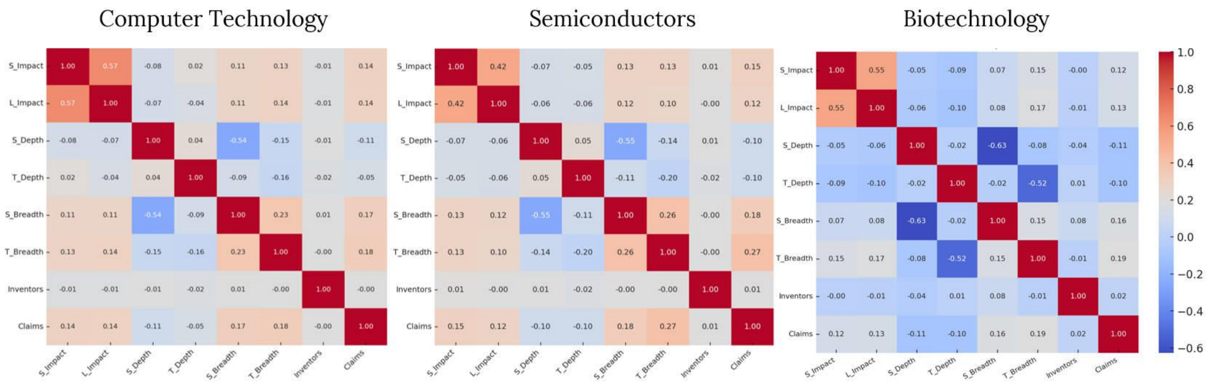

Figure 2. Heatmap of variable correlations in different fields.

As shown in Figure 2, the correlation coefficients among variables are generally low, all below 0.7. Although some relatively obvious correlations exist, most of the independent variables are within an acceptable range, not sufficient to trigger multicollinearity issues.

Models

Building on existing research, this section constructs negative binomial regression models to analyse the impact of various factors on patent impact. In terms of model setup, Models 0, 1, and 2 use the short-term impact of patents (S_Impact) as the dependent variable; Models 0`, 1`, and 2` use the long-term impact of patents (L_Impact) as the dependent variable. Models 0 and 0` incorporate only control variables to examine their independent impact on the dependent variable, thereby providing a baseline for comparison with subsequent models. Models 1 and 1` build on Models 0 and 0` by introducing independent variables, focusing on their direct effects on the dependent variable and preliminarily evaluating the research hypotheses. Models 2 and 2` further incorporate squared terms for the independent variables to test for nonlinear relationships between the independent and dependent variables, thereby more comprehensively revealing the complex internal connections between the variables.

To improve model stability and accuracy, the independent variables were normalized before modelling. For model evaluation, the alpha value measures the model's ability to account for overdispersion. Values close to 0 indicate low dispersion and a model closer to Poisson regression, while larger values indicate more pronounced overdispersion. Pseudo R^2 serves as a goodness-of-fit indicator; larger values indicate a greater ability of the model to explain the variation in the dependent variable. AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) are used to compare model goodness-of-fit and complexity; smaller values indicate a better balance between model fit and number of parameters. The following interpretations of the regression results in Table 4, Table 5, and Table 6 are presented regarding to three fields respectively.

Variable	Model 0	Model 1	Model 2	Model 0`	Model 1`	Model 2`
Intercept	-45.3627*** (0.2171)	-23.3046*** (0.2212)	-17.7338*** (0.2235)	-29.6927*** (0.2313)	-1.0009 (0.2354)	3.8320 (0.2357)
S_Breadth		0.1097*** (0.0010)	0.0683*** (0.0018)		0.1411*** (0.0010)	0.0342* (0.0019)
S_Depth		-0.0337*** (0.0005)	-0.1336*** (0.0013)		-0.0370*** (0.0006)	-0.2252*** (0.0014)
T_Breadth		0.1401*** (0.0005)	0.2225*** (0.0009)		0.1990*** (0.0006)	0.2975*** (0.0010)
T_Depth		0.0883*** (0.0005)	0.0893*** (0.0005)		-0.0084 (0.0006)	-0.0034 (0.0006)
$S\_Breadth^2$			0.0015 (0.0012)			0.0212* (0.0013)
S_Depth^2			0.0423*** (0.0004)			0.0759*** (0.0005)
$T\_Breadth^2$			-0.0643*** (0.0005)			-0.0772*** (0.0006)
T_Depth^2			-0.0327*** (0.0006)			-0.0255*** (0.0007)
Year	Yes	Yes	Yes	Yes	Yes	Yes
Inventors	-0.1287*** (0.00310)	-0.1120*** (0.00306)	-0.1061*** (0.00306)	-0.1196*** (0.00336)	-0.1069*** (0.00331)	-0.1027*** (0.00330)
Claims	0.0108*** (0.00003)	0.0095*** (0.00003)	0.0095*** (0.00003)	0.0135*** (0.00004)	0.0115*** (0.00003)	0.0114*** (0.00003)
alpha	1.1703*** (0.0007)	1.1404*** (0.0007)	1.1331*** (0.0007)	1.4729*** (0.0008)	1.4253*** (0.0008)	1.4135*** (0.0008)
Pseudo R^2	0.0058	0.0095	0.0104	0.0045	0.0090	0.0101
AIC	371,493	370,102	369,767	485,286	483,124	482,583
BIC	371,538	370,182	369,882	485,330	483,204	482,699

Note: ***, **, * indicate significance levels of $P < 0.01$, $P < 0.05$, and $P < 0.1$, respectively; standard errors are in brackets

Table 4. Models in computer technology.

Comparing the model fit indices reported in Table 4, Model 2 and Model 2`, which incorporate quadratic terms, provide the best fit for both short-term and long-term patent impact, as evidenced by the highest pseudo-R_2 and the lowest AIC and BIC values. Accordingly, the following analysis focuses on these two models. Regarding scientific knowledge breadth, S_Breadth shows a significantly positive effect on both short-term and long-term patent impact (S_Breadth: $\beta = 0.0683$, $p < 0.01$ in Model 2; $\beta = 0.0342$, $p < 0.05$ in Model 2`). In contrast, S_Depth exhibits a negative linear effect in the basic specification (S_Depth: $\beta = -0.0337$, $p < 0.01$ in Model 1), while the inclusion of the quadratic term reveals a significant U-shaped relationship (S_Depth^2: $\beta = 0.0423$, $p < 0.01$ in Model 2; $\beta = 0.0759$, $p < 0.01$ in Model 2`).For technological knowledge convergence, T_Breadth presents a significantly positive linear term and a significantly negative quadratic term (T_Breadth: $\beta = 0.2225$, $p < 0.01$; $T\_Breadth^2$: $\beta = -0.0643$, $p < 0.01$ in Model 2), indicating an inverted U-shaped relationship with patent impact. Similarly, T_Depth exhibits an inverted U-shaped relationship with short-term patent impact (T_Depth: $\beta = 0.0893$, $p < 0.01$; T_Depth^2: $\beta = -0.0327$, $p < 0.01$ in Model 2), whereas its long-term linear effect becomes statistically insignificant, while the quadratic term remains significantly negative (Model 2`), indicating diminishing returns over time.

Variable	Model 0	Model 1	Model 2	Model 0`	Model 1`	Model 2`
Intercept	-65.7025*** (0.2247)	-58.7938*** (0.2405)	-54.6107*** (0.2453)	-47.6266*** (0.2739)	-37.6453*** (0.2856)	-34.1385*** (0.2876)
S_Breadth		0.1752*** (0.0018)	0.3266*** (0.0027)		0.3018*** (0.0021)	0.4324*** (0.0031)
S_Depth		-0.0159 (0.0010)	0.1107*** (0.0020)		-0.0008 (0.0011)	0.0418* (0.0023)
T_Breadth		0.0433*** (0.0008)	0.1157*** (0.0012)		0.0675*** (0.0009)	0.1110*** (0.0014)
T_Depth		-0.0415*** (0.0007)	-0.0474*** (0.0007)		-0.0732*** (0.0009)	-0.0783*** (0.0009)
$S\_Breadth^2$			-0.0956*** (0.0019)			-0.1686*** (0.0022)
S_Depth^2			-0.0421*** (0.0007)			0.0073 (0.0008)
$T\_Breadth^2$			-0.0597*** (0.0007)			-0.0425*** (0.0008)
T_Depth^2			0.0014 (0.0008)			-0.0503*** (0.0010)
Year	Yes	Yes	Yes	Yes	Yes	Yes
Inventors	0.0339 (0.00348)	0.0353 (0.00348)	0.0272 (0.00348)	-0.0403 (0.00384)	-0.0297 (0.00383)	-0.0330 (0.00382)
Claims	0.0130*** (0.00006)	0.0108*** (0.00006)	0.0109*** (0.00006)	0.0162*** (0.00007)	0.0145*** (0.00007)	0.0145*** (0.00007)
alpha	1.0668*** (0.0010)	1.0529*** (0.0010)	1.0477*** (0.0010)	1.4986*** (0.0012)	1.4792*** (0.0012)	1.4722*** (0.0012)
Pseudo R^2	0.0110	0.0129	0.0136	0.0087	0.0102	0.0109
AIC	166,221	165,911	165,795	213,959	213,634	213,493
BIC	166,262	165,985	165,902	214,016	213,708	213,599

Table 5. Models in semiconductors.

The model results for the semiconductors are somewhat similar to those for the computer technology. Table 5 shows that Model 2 and Model 2` still provide the best fit in the semiconductors field. Regarding knowledge breadth, both scientific knowledge breadth and technological knowledge breadth exhibit significant inverted U-shaped relationships with patent impact at both time horizons (S_Breadth: $\beta = 0.3266$, $p < 0.01$; $S\_Breadth^2$: $\beta = -0.0956$, $p < 0.01$ in Model 2; $\beta = -0.1686$, $p < 0.01$ in Model 2`; T_Breadth: $\beta = 0.1157$, $p < 0.01$; $T\_Breadth^2$: $\beta = -0.0597$, $p < 0.01$ in Model 2; $\beta = -0.0425$, $p < 0.01$ in Model 2`). This indicates that moderately expanding the scope of knowledge coverage enhances patent impact, whereas excessive breadth weakens innovative performance. With respect to knowledge depth, scientific knowledge depth shows an inverted U-shaped relationship with short-term patent impact (S_Depth^2: $\beta = -0.0421$, $p < 0.01$ in Model 2), while its long-term linear effect becomes positive and statistically significant (S_Depth: $\beta = 0.0418$, $p < 0.05$ in Model 2`), suggesting that deeper scientific accumulation contributes to sustained patent influence over time. In contrast, technological knowledge depth exhibits a U-shaped relationship with short-term patent impact (T_Depth: $\beta = -0.0474$, $p < 0.01$; T_Depth^2: $\beta = 0.0014$, n.s. in Model 2), whereas its long-term effect turns significantly negative when accounting for nonlinearity (T_Depth^2: $\beta = -0.0503$, $p < 0.01$ in Model 2`), indicating diminishing returns from excessive technical specialization in the long run.

Variable	Model 0	Model 1	Model 2	Model 0`	Model 1`	Model 2`
Intercept	-5.3997* (0.2927)	25.5099*** (0.3126)	28.4864*** (0.3163)	-21.0596*** (0.2977)	20.5396*** (0.3097)	25.6915*** (0.3112)
S_Breadth		0.0679*** (0.0009)	0.1841*** (0.0027)		0.0920*** (0.0009)	0.3523*** (0.0027)
S_Depth		-0.0633*** (0.0015)	-0.0079 (0.0022)		-0.1072*** (0.0015)	0.0492** (0.0022)
T_Breadth		0.2334*** (0.0012)	0.2740*** (0.0015)		0.3094*** (0.0012)	0.3783*** (0.0016)
T_Depth		-0.0385*** (0.0008)	-0.0283*** (0.0009)		-0.0589*** (0.0008)	-0.0457*** (0.0009)
$S\_Breadth^2$			-0.0585*** (0.0013)			-0.1224*** (0.0013)
S_Depth^2			-0.0084 (0.0009)			-0.0509*** (0.0009)
$T\_Breadth^2$			-0.0327*** (0.0010)			-0.0487*** (0.0011)
T_Depth^2			-0.0198** (0.0009)			-0.0447*** (0.0009)
Year	Yes	Yes	Yes	Yes	Yes	Yes
Inventors	-0.0872*** (0.00244)	-0.0377 (0.00240)	-0.0365 (0.00239)	-0.0872*** (0.00244)	-0.0929*** (0.00239)	-0.0916*** (0.00239)
Claims	0.0132*** (0.00005)	0.0100*** (0.00005)	0.0100*** (0.00005)	0.0177*** (0.00005)	0.0136*** (0.00005)	0.0135*** (0.00005)
alpha	1.5516*** (0.0015)	1.4929*** (0.0015)	1.4905*** (0.0015)	1.8203*** (0.0013)	1.7400*** (0.0013)	1.7327*** (0.0012)
Pseudo R^2	0.0054	0.0113	0.0116	0.0058	0.0126	0.0133
AIC	161,653	160,701	160,668	265,930	264,119	263,957
BIC	161,695	160,777	160,777	265,972	264,195	264,066

Table 6. Models in biotechnology.

In the field of biotechnology, the model results exhibit more complex and unique dynamic characteristics. First, consistent with the previous two fields, Model 2 and Model 2` provide the best fit. Regarding knowledge breadth, both scientific knowledge breadth and technological knowledge breadth display significant inverted U-shaped relationships with patent impact at both time horizons (S_Breadth: $\beta = 0.1841$, $p < 0.01$; $S\_Breadth^2$: $\beta = -0.0585$, $p < 0.01$ in Model 2; $\beta = -0.1224$, $p < 0.01$ in Model 2`; T_Breadth: $\beta = 0.2740$, $p < 0.01$; $T\_Breadth^2$: $\beta = -0.0327$, $p < 0.01$ in Model 2; $\beta = -0.0487$, $p < 0.01$ in Model 2`). This indicates that broad knowledge integration promotes patent impact in biotechnology, while excessive diversification may weaken innovation performance. With respect to knowledge depth, scientific knowledge depth exhibits a significant inverted U-shaped relationship with long-term patent impact (S_Depth^2: $\beta = -0.0509$, $p < 0.01$ in Model 2`), whereas its short-term effect remains negative. In contrast, technological knowledge depth shows consistently negative effects on both short-term and long-term patent impact (T_Depth: $\beta = -0.0283$, $p < 0.01$ in Model 2; $\beta = -0.0457$, $p < 0.01$ in Model 2`), indicating that excessive technical specialization may hinder knowledge diffusion and interdisciplinary technology transfer in the biotechnology.

Taken together, the regression results across the three fields reveal clear and systematic differences between short-term and long-term patent impact. Across all three technological domains, knowledge breadth tends to play a more prominent role in shaping short-term outcomes, particularly in computer technology and semiconductors, where broader integration of scientific and technological knowledge is consistently associated with faster diffusion, higher early visibility, and stronger initial citation performance. This pattern suggests that, in fast-evolving technological environments, diversified knowledge inputs facilitate early-stage recognition and cross-domain dissemination of patent innovations. In contrast, the role of knowledge depth—especially scientific knowledge depth—becomes more differentiated when examining long-term patent impact. While deeper scientific foundations contribute positively to sustained influence in some fields, such as computer technology and semiconductors, their effects are nonlinear and vary substantially across technological contexts. In biotechnology, for example, excessive depth is more likely to constrain long-term diffusion, reflecting the field's strong dependence on interdisciplinary integration rather than narrow specialization.

Overall, these findings indicate that the structure of knowledge convergence affects not only the magnitude of patent impact but also its temporal profile. Knowledge breadth primarily accelerates short-term diffusion, whereas knowledge depth shapes the persistence and durability of influence over longer horizons.

Discussion and conclusions

This study investigates how knowledge convergence characteristics shape the short- and long-term impact of patents across different technological fields. By constructing multidimensional indicators of scientific and technological knowledge breadth and depth, the analysis reveals systematic differences in how knowledge structures influence patent impact over time.

The empirical results demonstrate that knowledge convergence exerts heterogeneous effects across time horizons. In the short term, knowledge breadth—particularly technological knowledge breadth—plays a more prominent role in facilitating patent visibility and early diffusion, often exhibiting either positive or inverted U-shaped effects across the three fields. This suggests that integrating diverse technological knowledge sources helps patents gain early recognition and accelerates initial dissemination. In contrast, in the long-term models, the role of knowledge depth becomes more pronounced. Scientific knowledge depth, especially in computer technology, exhibits significant nonlinear effects, indicating that sustained patent influence increasingly depends on the accumulation and integration of foundational scientific knowledge.

Cross-field comparisons further reveal distinct convergence patterns. Biotechnology patents show relatively stronger long-term responses to knowledge breadth, reflecting the field's interdisciplinary nature and its reliance on the integration of diverse scientific and technological domains. In contrast, the semiconductors exhibit more pronounced negative effects when knowledge depth becomes excessive, suggesting that over-specialization may constrain long-term diffusion and that innovation in this field may benefit more from focused breakthroughs along specific technological trajectories. Although the overall direction of the effects of knowledge convergence on patent impact is broadly consistent across fields, the magnitude and nonlinear forms of these effects vary substantially, highlighting the importance of aligning convergence strategies with both technological context and time scale.

Despite these contributions, this study has several limitations. First, the analysis relies primarily on U.S. patent data from the PatentsView platform, which may limit the generalizability of the findings to other institutional or regional contexts. Second, knowledge convergence is measured mainly through citation-based indicators, without incorporating additional information such as patent text semantics or inventor collaboration networks. Future research could be extended by integrating natural language processing and deep learning techniques to extract richer semantic features from patent texts, thereby improving the measurement of knowledge convergence. Moreover, patent impact needs not to be captured solely through citation counts. Graph-based and graph neural network approaches could be employed to jointly model patent attributes, convergence characteristics, and network topology, enabling dynamic predictions of patent influence and the identification of key patents across different time horizons.

Acknowledgements

We acknowledge support from the Shaanxi Innovation Capacity Support Program, Soft Science General Project (2024ZC-YBXM-073) and the National Science Foundation of China (71804114).

About the authors

Jing Ma is an Associate Professor in the School of Economics and Management, Xidian University, China. She received her PhD from Beijing Institute of Technology, and her research interests include technology forecasting, technology evaluation, and scientometrics. She can be contacted at majing@xidian.edu.cn.

Liyuan Li is a master student enrolled in the School of Economics and Management, Xidian University, China. Her research interests include knowledge linkage, patent bibliometrics, and technology forecasting. She can be contacted at 24061212955@stu.xidian.edu.cn.

Yaohui Pan is a Lecturer in the College of Economics and Management, China Jiliang University, China. She received her PhD from Beijing Institute of Technology, and her research interests include complex network analysis and knowledge management.

References

Battke, B., Schmidt, T. S., Stollenwerk, S., & Hoffmann, V. H. (2016). Internal or external spillovers—Which kind of knowledge is more likely to flow within or across technologies. Research Policy, 45(1), 27–41. https://doi.org/10.1016/j.respol.2015.06.014

Bhatt, P. C., Lai, K. K., Drave, V. A., & Sarkar, S. (2023). Patent analysis-based technology innovation assessment with the lens of disruptive innovation theory: A case of blockchain technological trajectories. Technological Forecasting and Social Change, 196, 122864. https://doi.org/10.1016/j.techfore.2023.122864

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. https://doi.org/10.1016/j.joi.2016.04.018

Choi, J., & Yoon, J. (2022). Measuring knowledge exploration distance at the patent level: Application of network embedding and citation analysis. Journal of Informetrics, 16(2), 101286. https://doi.org/10.1016/j.joi.2022.101286

Fernández, A. M., Ferrándiz, E., & Medina, J. (2022). The diffusion of energy technologies: Evidence from renewable, fossil, and nuclear energy patents. Technological Forecasting and Social Change, 178, 121566. https://doi.org/10.1016/j.techfore.2022.121566

Gan, J., & Zhang, Y. J. (2025). Dynamic analysis of patent value evolution: A study on the roles of novelty and impact. Technology Analysis & Strategic Management, 1–32. https://doi.org/10.1080/09537325.2025.2505077

Gao, Q., & Jiang, M. (2024). Exploring technology fusion by combining latent Dirichlet allocation with Doc2vec: A case of digital medicine and machine learning. Scientometrics, 129(7), 4043–4070. https://doi.org/10.1007/s11192-024-05069-1

Haessler, P., Giones, F., & Brem, A. (2023). The who and how of commercializing emerging technologies: A technology-focused review. Technovation, 121, 102637. https://doi.org/10.1016/j.technovation.2022.102637

Hong, S., Kim, J., Woo, H. G., Kim, Y. C., & Lee, C. (2022). Screening ideas in the early stages of technology development: A word2vec and convolutional neural network approach. Technovation, 112, 102407. https://doi.org/10.1016/j.technovation.2021.102407

Hou, J., Tang, S., Zhang, Y., & Song, H. (2023). Does prior knowledge affect patent technology diffusion? A semantic-based patent citation contribution analysis. Journal of Informetrics, 17(2), 101393. https://doi.org/10.1016/j.joi.2023.101393

Hu, Z., Zhou, X., & Lin, A. (2023). Evaluation and identification of potential high-value patents in the field of integrated circuits using a multidimensional patent indicator pre-screening strategy and machine learning approaches. Journal of Informetrics, 17(2), 101406. https://doi.org/10.1016/j.joi.2023.101406

Huang, Y., Li, R., Zou, F., Jiang, L., Porter, A. L., & Zhang, L. (2022). Technology life cycle analysis: From the dynamic perspective of patent citation networks. Technological Forecasting and Social Change, 181, 121760. https://doi.org/10.1016/j.techfore.2022.121760

Ke, Q. (2020). Technological impact of biomedical research: The role of basicness and novelty. Research Policy, 49(7), 104071. https://doi.org/10.1016/j.respol.2020.104071

Kong, J., Zhang, J., Deng, S., & Kang, L. (2023). Knowledge convergence of science and technology in patent inventions. Journal of Informetrics, 17(3), 101435. https://doi.org/10.1016/j.joi.2023.101435

Lai, I. C., & Su, H. N. (2024). Knowledge spectrum explored: Understanding source-recipient interactions and their influence on technology convergence. Technovation, 133, 103000. https://doi.org/10.1016/j.technovation.2024.103000

Li, L. M., Liu, Y. H., & Zhang, Y. F. (2024). The U-shaped relationship between patent age and patent value: Evidence from assigned patents in China. Economics of Innovation and New Technology, 33(6), 830–852. https://doi.org/10.1080/10438599.2023.2255978

Li, X., & Wang, Y. (2024). A novel integrated approach for quantifying the convergence of disruptive technologies from science to technology. Technological Forecasting and Social Change, 209, 123825. https://doi.org/10.1016/j.techfore.2024.123825

Liu, N., Mao, J., & Guan, J. (2020). Knowledge convergence and organizational innovation: The moderating role of relational embeddedness. Scientometrics, 125(3), 1899–1921. https://doi.org/10.1007/s11192-020-03684-2

Marx, M., & Fuegi, A. (2020). Reliance on science: Worldwide front-page patent citations to scientific articles. Strategic Management Journal, 41(9), 1572–1594. https://doi.org/10.1002/smj.3145

Marx, M., & Fuegi, A. (2022). Reliance on science by inventors: Hybrid extraction of in-text patent-to-article citations. Journal of Economics & Management Strategy, 31(2), 369–392. https://doi.org/10.1111/jems.12455

Moorthy, S., & Polley, D. E. (2010). Technological knowledge breadth and depth: Performance impacts. Journal of Knowledge Management, 14(3), 359–377. https://doi.org/10.1108/13673271011050102

Nakamura, H., Suzuki, S., Sakata, I., & Kajikawa, Y. (2015). Knowledge combination modeling: The measurement of knowledge similarity between different technological domains. Technological Forecasting and Social Change, 94, 187–201. https://doi.org/10.1016/j.techfore.2014.09.009

Oh, M., Jang, H., Kim, S., & Yoon, B. (2023). Main path analysis for technological development using SAO structure and DEMATEL based on keyword causality. Scientometrics, 128(4), 2079–2104. https://doi.org/10.1007/s11192-023-04652-2

Pan, W., Jian, L., & Liu, T. (2024). Knowledge generation and diffusion in science and technology: An empirical study of SiC-MOSFET based on scientific papers and patents. Technology Analysis & Strategic Management, 36(7), 1587–1603. https://doi.org/10.1080/09537325.2022.2106419

Szczygielski, K., & Mycielski, J. (2024). The mutual reinforcement of scientific and technological knowledge: A technology-level analysis. Scientometrics, 129(11), 6533–6549. https://doi.org/10.1007/s11192-024-05162-5

Wang, W., Wang, Z., Yu, T., Pak, C., & Yu, G. (2020). Research on citation mention times and contributions using a neural network. Scientometrics, 125(3), 2383–2400. https://doi.org/10.1007/s11192-020-03711-2

Wang, X., Zhao, Y., Liu, R., & Zhang, J. (2013). Knowledge-transfer analysis based on co-citation clustering. Scientometrics, 97(3), 859–869. https://doi.org/10.1007/s11192-013-1077-6

Xu, Y., Chen, X., Mao, J., & Li, G. (2025). Will patents with more interdisciplinary scientific knowledge have higher technological impact? Empirical evidence from USPTO patents. Scientometrics, 1–32. https://doi.org/10.1007/s11192-025-05280-8

Zhang, Y., Qian, Y., Huang, Y., Guo, Y., Zhang, G., & Lu, J. (2017). An entropy-based indicator system for measuring the potential of patents in technological innovation: Rejecting moderation. Scientometrics, 111(3), 1925–1946. https://doi.org/10.1007/s11192-017-2337-7