systems

MDPI

Article

Knowledge Recombination Reveals the Nonlinear Influence of Team Scale on Technological Breakthroughs

Le Song [1,*], Shan Chen [1], Jinqiao Liang [1] and Xiao Yin [2]

[1] Business School, Shandong Normal University, Jinan 250014, China; chens202508@163.com (S.C.); 18565450647@163.com (J.L.)
[2] School of Business Administration, South China University of Technology, Guangzhou 510641, China; yx16_wk@163.com
* Correspondence: songle@sdnu.edu.cn

Abstract

In the knowledge economy era, optimizing R&D team size is crucial for breakthrough innovation. Breakthrough technologies rely more on knowledge restructuring and technological leaps than general technologies do. However, it remains unclear whether breakthrough technology formation follows a simple "more people, more power" logic within technological systems. This work examines 35,955 patents in recommendation system technology to propose a relationship model between collaboration scale and breakthrough technological innovation based on patent data from the recommendation system field. It aims to elucidate how collaboration scale influences breakthrough technological innovation through knowledge restructuring, thereby providing theoretical support and practical guidance for enterprises, institutions, and governments in innovation activities to advance technological innovation. The findings reveal three key points: (1) The relationship between collaboration scale and breakthrough innovation is not linear but follows an inverted U-shaped curve; (2) Knowledge recombination significantly mediates this relationship, also exhibiting an inverted U-shaped pattern with collaboration scale; (3) The inverted U-shaped effect of collaboration scale on breakthrough innovation varies by country. The optimal thresholds are 14.058 entities for China, 57.151 entities for the United States, and 4.801 entities for Russia. This work breaks through the limitations of the traditional theoretical framework and constructs a three-dimensional analysis framework of "collaboration scale → knowledge recombination → breakthrough technological innovation". By introducing the mediating variable of knowledge recombination, this paper reveals the mechanism of R&D team size on radical innovation. It provides a theoretical basis for the construction of an innovation team and provides a theoretical basis for enterprises, governments, and institutions.

Keywords: technology system; breakthrough technological innovation; collaboration scale; knowledge recombination; recommendation system

check for updates

Academic Editor: Mitsuru Kodama

Received: 1 September 2025
Revised: 3 October 2025
Accepted: 5 October 2025
Published: 7 October 2025

Citation: Song, L.; Chen, S.; Liang, J.; Yin, X. Knowledge Recombination Reveals the Nonlinear Influence of Team Scale on Technological Breakthroughs. *Systems* **2025**, *13*, 877. https://doi.org/10.3390/systems13100877

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

As an indispensable part of human development, the innovation and iteration of technological systems are the main driving forces promoting social progress and economic growth [1]. Breakthrough technological innovation refers to innovations based on new scientific, technological, and engineering knowledge, which can completely change the technological foundation and market structure of enterprises [2]. It has two forms of expression: jump-type technology enhancement that follows the existing technology trajectory

and new technology trajectory transition development that breaks the existing technology system [3,4]. However, the realization of breakthrough innovation is often accompanied by high risk and high investment [5]. Governments and researchers are still actively exploring new ways to achieve breakthrough technological innovation.

Knowledge creation and technological innovation have become the most important links in market competition [6]. As a key link in the process of innovation, knowledge recombination can effectively apply existing knowledge resources by integrating professional knowledge in different fields so that new knowledge and innovation achievements can be continuously produced [7]. At the same time, knowledge recombination can promote the flow and sharing of knowledge among different subjects, such as enterprises, universities, and research institutions; help to build a benign innovation ecosystem; and ultimately promote the overall efficiency of the technology system. Because breakthrough technological innovation is a leap-forward technology enhancement that follows the existing technology trajectory, or a new technology leap that breaks the existing technology system, it is more dependent on knowledge recombination [8]. Given the unique characteristics of the knowledge resource structure in the technology system and the new knowledge needs, the cross-enterprise and cross-institutional knowledge resource organization system [9] and the utilization rate of knowledge resources [10] have become an important way to achieve breakthrough technological innovation in the technology system.

Given the importance of knowledge recombination in the context of the vigorous development of digital technology, cooperation has become an important form for individuals or organizations to carry out innovation activities [11]. Teamwork enables the author to divide the task. On the premise that the author and his collaborators have professional knowledge, respectively, the research team integrates the collaborators' professional knowledge, creativity, and skills, and at the same time organically matches or reconstructs these integrated contents with their professional knowledge [12]. In recent years, the collaboration scale has had an important influence on breakthrough technological innovation, and knowledge restructuring has played a key role as a bridge between the collaboration scale and breakthrough technological innovation. The collaboration and cooperation of technology R&D personnel is the key to improving the overall efficiency of the technology system. "More people, more power" is an important organizational pattern for technology research [13]. Breakthrough technological innovation cooperation is not more complex general innovation, but a paradigm revolution from transactional collaboration to evolutionary symbiosis to enhance cooperation resilience to confront stronger uncertainty in the R&D process [14]. It is still unknown whether the R&D pattern of "more people, more power" is suitable for breakthrough technological innovation.

Scholars have comprehensively examined and analyzed R&D cooperation, breakthrough technological innovation, and their relationship, providing a valuable reference for promoting the realization of breakthrough technological innovation. However, most of them remain at the level of theoretical framework and method, and there are few relevant empirical studies, lacking the combination and verification of theoretical research and practical cases. Although existing research has explored the relationship between R&D cooperation networks and breakthrough technological innovation, systematic research on how collaboration scale affects breakthrough technological innovation through knowledge recombination is scarce. Most of the existing literature focuses on a single level of analysis, lacking in-depth discussion of the complex mechanisms in the innovation process. Therefore, there are still many challenges for exploration on the mechanism between the collaboration scale and breakthrough technological innovation. As a key driving factor of breakthrough technological innovation, the mediating effect of knowledge recombination in the process of cooperation amplification has not been fully discussed. This work uses

patent data in the field of recommendation systems to construct a relationship model between "collaboration scale → knowledge recombination → breakthrough technological innovation", aiming to clarify how collaboration scale affects the formation of breakthrough technological innovation through knowledge recombination, to provide theoretical support and practical guidance for enterprises, institutions, and governments in innovation activities, andto promote the progress of technological innovation. The main contributions of this work are as follows: (1) The relationship between collaboration scale and breakthrough technological innovation is empirically analyzed, which enriches the theoretical system of cooperation synergy and breakthrough technological innovation. (2) It reveals the nonlinear influence of collaboration scale on radical technological innovation from a new perspective of knowledge recombination. (3) The scale of cooperation is deconstructed from the dual dimensions of the number of inventors and the number of institutions to avoid the single influence of the linear scale. (4) The heterogeneity of collaboration scale and breakthrough technological innovation in different countries and regions is analyzed, which provides a realistic basis for the government to introduce relevant regional policies to promote the development of breakthrough technological innovation.

The rest of the chapters are arranged as follows: Section 2 combines the theoretical basis and demonstrates the research hypothesis; empirical data, variable measures, and fitting models are introduced in Section 3; Section 4 gives the statistical description of the empirical data, tests the research hypothesis, and further tests the endogeneity, robustness, and heterogeneity. Section 5 summarizes the research, management enlightenment, and future work.

2. Theoretical Analysis and Research Hypothesis

2.1. Theoretical Basis

Knowledge recombination theory is an important theoretical framework to explain technological innovation, especially breakthrough technological innovation [8]. With the deepening of innovation research, scholars generally believe that the essence of innovation lies in the reintroduction and integration of existing knowledge elements, which makes knowledge recombination a key source of innovation. The theoretical context began with Schumpeter's pioneering work, which pointed out that the enterprise is a collection of knowledge elements, and innovation is derived from a new combination of these knowledge elements [15]. Follow-up research continues to deepen along this core idea. For example, some studies define knowledge recombination as the process of combining objective material elements to create new materials under the existing technical level and emphasize that knowledge innovation is essentially a reintegration of existing knowledge [16]. Furthermore, the research points out that most of the innovation achievements in reality can be attributed to the recombination of existing concepts or material objects [17], and even some studies have found that the creation of new things in other fields mainly comes from the recombination of existing things [18]. These theoretical viewpoints have been supported in empirical research. For example, empirical analysis of the semiconductor industry has confirmed that restructuring new knowledge outside the enterprise can significantly improve its innovation ability [19]. On the whole, knowledge recombination can be defined as a dynamic process in which enterprises explore, integrate, and reconstruct various knowledge elements in the internal and external environment and then produce new results.

The core attributes of knowledge recombination capability can be summarized into three dimensions: diversity [20], synergy [21], and knowledge span [22]. Diversity represents the abundance and type distribution of heterogeneous knowledge elements in the enterprise knowledge system. Driven by the digital wave, the iterative upgrading of

information technology has significantly broadened the breadth of enterprise knowledge reserves and catalyzed unprecedented technological integration possibilities. Cooperativity reflects the internal correlation strength between knowledge units, and the coupling degree of knowledge elements directly affects the feasibility of recombination operation. The plasticity of digital platforms effectively reduces the barriers to cross-domain technology integration and creates favorable conditions for knowledge collaboration. The knowledge span is used to measure the technological generation difference between different knowledge systems. The frequent cross-organizational knowledge interaction in the digitization process not only causes cognitive conflicts but also promotes breakthrough technological innovation and enables enterprises to realize technological paradigm transformation.

2.2. The Influence of Collaboration Scale on Breakthrough Technological Innovation

Inventors are the implementers and finalists of innovation activities, and their cooperative R&D network has an important influence on their ability to obtain resources and achieve innovation breakthroughs. At present, due to the increasing demand for technology, cross-domain technology is more complex and the required knowledge is more extensive, which will inevitably prompt technology R&D personnel to actively seek cooperation. At the same time, with the development of communication technology, transportation is more convenient, the limitation of geographical dimension is reduced, and the scale of team cooperation is easier to expand [23,24].

Team cooperation provides human resources and intellectual support for technical research and forms a cohesive force of "more people, more power" to a certain extent [11]. In the process of breakthrough technological innovation, do more people mean more power? This issue is still worthy of further exploration. When the scale of teamwork expands within a certain range, it will promote breakthrough technological innovation. On the one hand, the open innovation theory holds that R&D subjects actively carry out external cooperation activities in the face of a complex and changeable external environment to obtain more abundant external information and knowledge, to improve the efficiency of internal resource utilization and promote the output of innovation results [25]. On the other hand, the social capital theory [26] holds that the more social groups, organizations, and individuals contact, the more social resources they usually have. Correspondingly, their capital strength is more abundant, which is conducive to breakthrough technological innovation. However, when the scale of teamwork exceeds a certain extent, too large a collaboration scale is not conducive to breakthrough technological innovation. On the one hand, according to the transaction cost theory, when the team scale exceeds a certain range, the cost of management and coordination increases, which is not conducive to breakthrough technological innovation [27]. On the other hand, according to the attention allocation theory [28], due to the cooperation with external organizations, there is a need to pay a lot of energy to establish and maintain the relationship. Excessive collaboration scale will lead to distracted attention of inventors, which is not conducive to digging up valuable resources or ideas, thus inhibiting breakthrough technological innovation. Therefore, this work proposes Hypothesis 1 (H1).

Hypothesis 1 (H1). *There is an inverted U-shaped relationship between collaboration scale and breakthrough technological innovation, that is, the initial growth of collaboration scale can promote the emergence of breakthrough technological innovation, but beyond a certain scale, it will inhibit its formation.*

2.3. The Mechanism of Knowledge Recombination

Cooperation can achieve an innovation transition through knowledge recombination [29]. Therefore, in the process of breakthrough technological innovation, knowledge

recombination is the main goal of cooperation and the key driving factor of breakthrough technological innovation. On the one hand, the collaboration scale has an inverted U-shaped influence on knowledge recombination. In the face of complex technological research needs, cooperation is an important way of knowledge production and transfer [30]. The process of cooperation must involve the application, transformation, and recombination of knowledge in related fields. Inventors have different knowledge reserves, material resources, and R&D capabilities. Through cooperation, they can promote the recombination of team knowledge, promote technological transition and complementary advantages, and achieve "1 + 1 > 2" technology R&D synergy. Cooperation is of great significance to the realization of disciplinary leapfrogging and scientific breakthroughs [31]. In the early stage of the expansion of the collaboration scale, it is possible to improve the efficiency of restructuring in a non-redundant combination by introducing interdisciplinary heterogeneous knowledge elements, which is particularly evident in the field of complex technology [32]. However, excessive collaboration scale may also have negative effects, such as increasing the difficulty of knowledge identification, leading to the dislocation of technology supply and demand, and increasing transaction costs and information asymmetry [33]. With the expansion of the scope of cooperation, participants may tend to be similar in cognitive structure, research ideas, and knowledge base, forming homogeneous resource redundancy but inhibiting the recombination and creation of new knowledge [34]. At the same time, too many cooperative relationships will also increase the cost of coordination, reduce communication and trust among members, make knowledge recombination difficult to carry out effectively, and even lead to the failure of cooperation [35]. On the other hand, knowledge restructuring itself has a significant role in promoting the output of breakthrough innovation results. Based on the patent data of 26 years in the field of biotechnology in the United States [36], some studies have found that the combination of new components can not only improve the average utility of innovation but also significantly increase the probability of breakthrough innovation and reduce the risk of R&D failure. Similarly, the empirical analysis based on Chinese manufacturing listed companies also shows that there is a significant positive correlation between knowledge restructuring activities and breakthrough innovation performance [11]. Knowledge recombination is considered to be an important way for teams to gain innovation advantages in the era of big science [37]. It plays an important role in the process of achieving breakthrough innovation and is one of the key links to promoting breakthrough innovation. Therefore, this work proposes Hypothesis 2 (H2).

Hypothesis 2 (H2). *The inverted U-shaped effect of collaboration scale on breakthrough technological innovation stems from the mediating effect of knowledge recombination.*

3. Research Design

3.1. Data Sources

With the development of artificial intelligence, the recommendation system is in a critical window period of transition from deep learning to generative intelligence. There are not only continuous collaborative filtering and deep representation learning patents, but also mutant patents that use diffusion models and reinforcement learning for sequence recommendation. This "continuous + mutation" technology pedigree provides technological opportunities for breakthrough technological innovation. Recommendation system patents are highly dependent on the integration of computer science, knowledge engineering, and cross-domain knowledge. They have both technical heterogeneity, cooperation diversity, and data integrity, and are more representative of breakthrough technological innovation. Therefore, this work focuses on the related technical fields of the recommenda-

tion system, collects multi-source data, and conducts empirical analysis. The data mainly comes from the following platforms:

(1) Patent search platform PatentGuru https://www.patentguru.com/cn (accessed on 11 March 2024). The patent retrieval platform has more than 170 million patent records in 175 countries and regions. It can optimize the search results according to the user-defined search formula, combined with natural language processing, big data cloud processing, and machine learning technology. It does not require users to preprocess the search words to ensure the comprehensiveness and accuracy of the search. Using the combination of domain name and key research content, the patent search formula of the PatentGuru platform is set for this field: (tiab = (Recommendation Systems) OR tiab = (Collaborative Filtering) OR tiab = (Content Filtering) OR tiab = (Hybrid Filtering) OR tiab = (Rule-based Recommendation) OR tiab = (Ensemble Learning) OR tiab = (Context-aware Recommendation) OR tiab = (Exploration and Exploitation) OR tiab = (Reinforcement Learning) OR tiab = (Timing recommendation)) AND (fd = 1990:2023). In this work, 45,665 patents were searched by using this search term, and the basic description items such as title, abstract, publication (announcement) number, application date, IPC classification number, and so on were retained. Further excluding the missing data of title, abstract, public (announcement) number, public (announcement) date, and IPC classification number, and the data of Hong Kong, Macao, and Taiwan, 37,243 patents were retained.

(2) Patent search platform Incopat https://www.incopat.com/ (accessed on 11 March 2024). The patent retrieval platform has a set of independent research and development of a mature patent value evaluation system. This evaluation system selects the commonly used patent value evaluation indicators in the industry and forms the evaluation values of technical stability, technological advancement, and shared value, which provides a reference for users to quickly browse or select key patents. This work uses this platform to obtain the technical stability and shared value of the above 37,243 patents.

(3) World Bank Open Data https://data.worldbank.org.cn/ (accessed on 26 March 2025). The platform provides a range of World Bank data sets, including databases, formatted tables, reports, and other resources. Provide data obtained through sampling surveys of households, business organizations, or other institutions. Explore the purchasing power parity, price level, economic data, and methods behind the world's most extensive statistical cooperation. This work collects the total GDP, per capita GDP, R&D expenditure, and other data of 58 countries and regions involved in the patent data on the platform.

After further processing the above three sets of data, this work removes some data with missing control variables and patent data that cannot be calculated for knowledge recombination in the first five years and finally retains 35,955 patent data in the field of recommendation systems for empirical analysis.

3.2. Variable Definition and Measure

3.2.1. Explained Variable: Breakthrough Technological Innovation (BTI)

Breakthrough technology includes not only the research direction that does not meet the mainstream demand but also the innovative application that changes the industrial trajectory. Its patent citation is one of the key indicators for evaluating breakthrough technology. This work draws on the research of Vakili and Kaplan [38], Fontana et al. [39], and uses the patent citation to measure the degree of breakthrough technological innovation, and the citation of patent i is recorded as BTI_i, due to the lag and cumulative effect of citations, to ensure the fairness of technology screening, this work adopts a dynamic evaluation perspective. The citation of patent i in the data collection year (2024) is recorded

as BTI_i, and the patents in the same application year are sorted and compared. The top 5% of the patents cited are considered to have breakthrough technological innovation, with a value of 1. Otherwise, the value is 0. The critical citation of each application year is recorded as $cite_{top5\%}$. Formula (1) is used to measure whether the patent i is a breakthrough technological innovation (BTI_i).

$$BTI = \begin{cases} 1, & cite_i \geq cite_{top5\%} \\ 0, & cite_i < cite_{top5\%} \end{cases} \quad (1)$$

3.2.2. Explanatory Variable: Collaboration Scale (CS)

Compared with the traditional indicators of collaboration scale at the organizational level, such as the number of enterprises and the scale of funds, this work refers to the micro- and meso-two-dimensional measurement of Gao [40], and uses the product of the number of inventors (N_{inv}) and the number of participating institutions (N_{ins}) to measure the collaboration scale of patent i (CS_i), which not only avoids the problem of dimensional inconsistency faced by factor addition but also avoids the one-sidedness of a single dimension. The specific measurement is shown in Formula (2).

$$CS_i = N_{inv} \times N_{ins} \quad (2)$$

3.2.3. Mediating Variable: Knowledge Recombination (KR)

When calculating the knowledge recombination level of the patent i applied for in the observation year t, the reference literature [8,38] sets a 5-year window period, that is, the $t-5$ to $t-1$ year period is used as a control to construct the benchmark knowledge base set_{win}, in which the four IPC classification numbers are used as the proxy elements of knowledge. $set_{win}^{(t-5)}$ is a binary combination set of all IPC classification numbers appearing in the same patent during the above window period. By comparing the IPC classification number set_i^t of the patent i applied for in the observation year t with the IPC classification number $set_{win}^{(t-5)}$ in the window period, all the IPC classification number $set_i^t - (set_i^t \cap set_{win}^{(t-5)})$ that are not in $set_{win}^{(t-5)}$. Using $|set_i^t - (set_i^t \cap set_{win}^{(t-5)})|$ to measure the knowledge recombination ability KR_i of the patent i applied by the observation year t, and using Formula (3) to measure. For example, it is assumed that there are only three technical combinations in the whole technical field from $t-5$ to $t-1$: the combination of D01N and F16B (D01N + F16B), the combination of D01N and C07M (D01N + C07M), and the combination of G06D and H04E (G06D + H04E). Then the benchmark library set_{win} is D01N + F16B, D01N + C07M, and G06D + H04E. Next, we analyze the patent i applied in the t year, and extract all IPC binary combinations contained in it to form a set_i^t. It is assumed that the patent i combines the two technologies of D01N and G06D, that is, set_i^t = D01N + F16B (old combination), D01N + G06D (new combination). Subsequently, we find a completely "novel" combination in patent i that has not appeared in the benchmark library through set operation, that is, $|set_i^t - (set_i^t \cap set_{win}^{(t-5)})|$. In this example, patent i contains an old combination (D01N + F16B) and a new combination (D01N + G06D). The intersection of the old combination and the benchmark library is not empty, while the new combination (D01N + G06D) does not exist in the benchmark library. Therefore, the novel combination set is D01N + G06D. Finally, the knowledge recombination ability of patent i (KR_i) is the number of elements in this novel set. In this case, |D01N + G06N| = 1. The larger the value is, the higher the degree of breakthrough knowledge recombination is, indicating that the patent breaks the existing technical path.

$$KR_i = |set_i^t - (set_i^t \cap set_{win}^{(t-5)})| \quad (3)$$

3.2.4. Control Variable

In order to eliminate the influence of other factors on the empirical analysis, the following control variables are selected in this work.

(1) Knowledge accumulation (KA) [41]: Knowledge accumulation refers to the accumulation of technology, experience, and information obtained by the organization in the long-term R&D process. A high level of knowledge accumulation can promote effective communication and collaboration between partners, thereby improving the efficiency and quality of innovation. Therefore, controlling knowledge accumulation can ensure that the relationship between the collaboration scale and breakthrough technological innovation in the research is not interfered with by the difference in the knowledge base, which makes the research results more reliable. In this work, the total number of IPC classification numbers $N^t_{(j,IPC)}$ of the patent's country and region j as of the observation year t is used to measure knowledge accumulation, using Formula (4).

$$KA_j = N^t_{(j,IPC)} \tag{4}$$

(2) Technical stability (TS) [42]: A stable technical foundation can reduce risks and increase the confidence of partners in innovative projects, which in turn affects the implementation process and results of innovation. By controlling the stability of technology, it can eliminate its interference with the relationship between the collaboration scale and breakthrough technological innovation and ensure the reliability of research results. This work uses the evaluation value of patent technology stability provided by the InCopat database. The evaluation value is obtained by evaluating the status of patent review, whether there is litigation, whether there is pledge preservation, and whether the application is invalid. The score is assigned on a scale of 1–10. The higher the number is, the more stable the technology is.

(3) GDP per capita (GDP) [43]: Countries and regions with higher GDP per capita usually have a more mature economic environment, rich R&D resources, and perfect policies and regulations. These factors may significantly promote the development of innovation activities, accelerate industrial upgrading and innovation, and promote the output of breakthrough innovation results. In order to eliminate the influence of dimension, this work takes the logarithm of GDP.

(4) R&D expenditure (RD) [44]: R&D expenditure is the investment of enterprises or institutions in technological innovation and product development, which directly affects the ability and results of innovation. Higher R&D expenditure usually means more resources are invested in innovation activities, which may lead to higher levels of breakthrough technological innovation. Therefore, taking R&D expenditure as a control variable can effectively eliminate the influence of capital investment on the relationship between collaboration scale and innovation results so as to more accurately analyze the role of collaboration scale.

(5) The region where the patent application is located (C) [45]: Control the variables of the continent where the patent application is located, eliminate the influence of environmental and resource differences on the research results, and divide the geographical location of the country and region where the patent application is located into North America, Oceania, Africa, South America, Europe, and Asia. In Stata 17, the encode command is used to assign these categories numerical values from 1 to 6, converting the character variable into a numeric variable.

(6) Shared value (SV) [46]: As a comprehensive indicator, it integrates multiple dimensions to evaluate the potential value of patents and uses it as a control variable to effectively strip the influence of patent value differences on breakthrough techno-

logical innovation. This work uses the patent sharing value evaluation provided by the InCopat database. The evaluation value is obtained by evaluating the number of R&D personnel, whether licensing or transfer, the number of claims, the number of family patents and other dimensions, and is assigned by 1–10. The higher the number, the more significant the market value and influence of the technology.

(7) Collaboration mode (TC) [47]: Regional collaboration models are categorized into international collaboration and internal collaboration. If Patent i adopts the international collaboration model, TC_i is recorded as 1; otherwise, TC_i is recorded as 0. The control variable is adjusted to eliminate the influence of the collaboration model difference.

3.3. Model Construction

As shown in Formula (1), the dependent variable (whether it is a breakthrough technological innovation patent) is a binary variable, with a value of 0 or 1, and it is asymmetric. The research needs to analyze the influence mechanism of explanatory variables on the probability of event occurrence. The Logit model assumes that the potential decision-making process obeys the logistic distribution, which can effectively deal with the nonlinear characteristics of the binary choice problem. Therefore, the Logit model is used as the model basis for empirical analysis.

To examine whether the collaboration scale has a nonlinear effect on breakthrough technological innovation, that is, to verify Hypothesis 1 (H1), Formula (5) is constructed for subsequent empirical analysis, where BTI_i denotes the dependent variable and CS_i denotes the independent variable, $Controls_i$ represents all control variables, and ξ_i is the error term. The formula of the specific model is as follows.

$$BTI_i = \alpha_0 + \alpha_1 CS_i + \alpha_2 CS_i^2 + \sum \alpha_k Controls_i + \xi_i, \quad (5)$$

where α_0 is the intercept term, and the coefficients of each variable are α_1, α_2, α_K.

In order to further study the mediating effect of knowledge recombination, Formulas (6), (7), and (8) are constructed, respectively.

$$KR_i = \omega_0 + \omega_1 CS_i + \omega_2 CS_i^2 + \sum \omega_m Controls_i + \xi_i, \quad (6)$$

$$BTI_i = \psi_0 + \psi_1 KR_i + \sum \psi_n Controls_i + \xi_i, \quad (7)$$

$$BTI_i = \beta_0 + \beta_1 CS_i + \beta_2 CS_i^2 + \beta_3 KR_i + \sum \beta p Controls_i + \xi_i, \quad (8)$$

where KR_i is the knowledge recombination level of patent i, ω_0, ψ_0 and β_0 are intercept terms, ω_1, ω_2, ω_m, psi_1, psi_n, β_1, β_2, β_3 and β_p are variable coefficients, and the content setting of the remaining Formulas is consistent with Formula (5).

4. Empirical Analysis

4.1. Descriptive Statistical Analysis

Based on the recommended system patent data collected in this work, the descriptive statistics of the research variables are shown in Tables 1–3. Among the statistical characteristics of each variable, the dependent variable is a binary variable representing breakthrough technological innovation. The value can only be 0 or 1. The number of patents with a value of 1 is 1964, and the number of patents with a value of 0 is 33,991, accounting for 94.5% of the total. From the perspective of collaboration scale, the average value is 4.381, and cooperation is still an important path for technology research and development. The standard deviation is 4.838, the minimum value is 1, and the maximum value is 169. It reflects that

there is a certain degree of imbalance in the collaboration scale. Large-scale cooperation exists, but the number is small.

To test the collinearity between variables, Table 4 shows the Pearson correlation coefficient between variables, and most variables do not have collinearity. The correlation coefficient between CS and CS^2 is 0.735, which is due to the high correlation between the first term and the second term. The correlation coefficient between technical stability (TS) and knowledge accumulation (KA) is 0.683, but the measurement of technical stability is based on the evaluation of qualitative indicators such as patent audit status and litigation behavior, and knowledge accumulation is based on the cumulative measurement of IPC classification numbers. They are heterogeneous variables. The correlation coefficient between GDP per capita and the applicant's region C is −0.633, but GDP per capita is an economic indicator. The applicant's region C is a dummy variable for classification, which is also a heterogeneous variable. In addition, the dummy variable is only used for classification when using Stata analysis. In view of the high Pearson correlation coefficient between the above three pairs of variables, the variance inflation factor (VIF) is further used to test the multicollinearity between variables, and the results are shown in Table 5. The VIF values of all variables are less than the standard value of 10 for judging multicollinearity, which proves that there is no multicollinearity problem in each variable.

Table 1. Descriptive statistical analysis table of related variables.

Variable	Mean	Standard Deviation	Minimum	Median	Maximum
CS	4.38	4.838	1	3	169
CS^2	42.59	296.994	1	9	28,561
KA	6227.10	5792.345	1	3838	18,685
TS	5.53	3.174	1	3	10
GDP	9.56	1.028	6.110	9.225	11.699
RD	30.31	1.269	20.357	30.691	32.039
SV	8.50	1.577	2	9	10

Table 2. Descriptive statistical analysis table of variables in the applicant's region.

	BTI			
	Frequency	Percentage	Effective Percentage	Cumulative Percentage
0	33,991	94.5%	94.5%	94.5%
1	1964	5.5%	5.5%	100.0%
total	35,955	100.0%	100.0%	-

Table 3. Descriptive statistical analysis table of variables in the applicant's region.

	C			
	Frequency	Percentage	Effective Percentage	Cumulative Percentage
1	6317	17.569	17.569	17.569
2	55	0.153	0.153	17.722
3	2	0.006	0.006	17.728
4	7	0.019	0.019	17.747
5	4370	12.154	12.154	29.9017
6	25,204	70.099	70.099	100.000
total	35,955	100.000	100.000	

Table 4. Correlation coefficient table of variables.

Variables	CS	CS^2	KA	TS	GDP	C	RD	SV
CS	1.000							
CS^2	0.735	1.000						
KA	0.165	0.047	1.000					
TS	0.173	0.048	0.683	1.000				
GDP	−0.127	−0.042	−0.251	−0.421	1.000			
C	0.076	0.020	0.301	0.363	−0.633	1.000		
RD	0.076	0.006	0.543	0.338	0.309	−0.321	1.000	
SV	0.029	0.008	0.051	0.254	0.347	−0.289	0.366	1.000

Table 5. VIF test.

Variable	VIF	SQRT VIF	Tolerance	R-Squared
CS	2.30	1.52	0.4349	0.5651
CS^2	2.22	1.49	0.4514	0.5486
KA	3.06	1.75	0.3273	0.6727
TS	2.90	1.70	0.3442	0.6558
GDP	2.34	1.53	0.4268	0.5732
C	2.12	1.46	0.4719	0.5281
RD	2.61	1.61	0.3838	0.6162
SV	1.62	1.27	0.6167	0.3833

4.2. Hypothesis Testing

Logit regression analysis is used to test the research hypothesis under the control of knowledge accumulation, applicant's region, technical stability, shared value, R&D investment, etc. Formula (5) shows the direct effect of collaboration scale on breakthrough technological innovation under the Logit model. Formula (6) presents the direct effect of the collaboration scale on the ability of knowledge recombination after OLS regression. Formula (7) is used to verify the direct effect of knowledge recombination on radical technological innovation. Based on Formula (5), Formula (8) introduces the variable of knowledge recombination to explore the mediating role of knowledge recombination between collaboration scale and breakthrough technological innovation.

4.2.1. The Nonlinear Effect of Collaboration Scale on Breakthrough Technological Innovation

According to the regression analysis results of Table 6, the pseudo-R^2 value in the Formula (6) model is 0.199, and the model has a high degree of fit. The model shows that the first-order coefficient of the collaboration scale is positive and significant at the 1% level, and the second-order term is negative and significant at the 5% level, initially supporting Hypothesis 1 (H1). To further verify the robustness of the "inverted U-shaped" relationship, this work uses the nonlinear relationship test method proposed by Lind and Mehlum [48], and uses the Utest statistic to test. The test results show that the slopes of the left and right endpoints of the curve are 0.064 and −0.128, respectively, which are significant at the 5% level. This indicates that the collaboration scale and breakthrough technological innovation show an inverted U-shaped relationship; that is, with the increase in the collaboration scale, the possibility of patents as breakthrough technological innovation increases first and then decreases, which is consistent with the hypothesis proposed in this paper. Therefore, Hypothesis 1 (H1) can be verified.

Table 6. Regression analysis results.

Variable	BTI Formula (5) H1	KR Formula (6)	BTI Formula (7)	BTI Formula (8) H2
CS	0.06510 *** (0.00000)	0.00433 ** (0.01269)		0.06482 *** (0.00000)
CS^2	−0.00057 ** (0.02428)	−0.00006 ** (0.01380)		−0.00056 ** (0.02548)
KR			0.04525 *** (0.00164)	0.04464 *** (0.00214)
Cons_	−14.62981 *** (0.00000)	2.10601 *** (0.00000)	−14.49335 *** (0.00000)	−14.80347 *** (0.00000)
Controls	YES	YES	YES	YES
N	35,955	35,955	35,955	35,955
adj. R^2		0.041		
Pseudo R^2	0.199		0.195	0.200

The robust standard error is in parentheses. **, *** are significant at the 5%, 1% level, respectively.

4.2.2. The Mediating Effect of Knowledge Recombination

The previous analysis shows that the collaboration scale has an inverted U-shaped effect on breakthrough technological innovation, which first promotes and then inhibits. In order to verify the mediating effect of knowledge recombination, the widely used Bootstrap method is first used to test the existence of the mediating effect of knowledge recombination. It should be pointed out in particular that the mediating path assumed in this study involves a nonlinear relationship (collaboration scale → breakthrough technological innovation is nonlinear). In this case, the indirect effect size of Bootstrap method may not fully capture the actual importance of the mediation process because the effect of the path may change at different levels of the predictor. Therefore, a numerically smaller average indirect effect may still mark the existence of a theoretical and nonlinear intermediary channel. Although the point estimate of the indirect effect is small, its statistical significance indicates that knowledge recombination does play a mediating role between the collaboration scale and breakthrough technological innovation. This finding supports our hypothesis about nonlinear mediation. The specific discussion is as follows.

The parameter IND is generated, and the significance and strength of IND are used to judge whether there is a mediating effect. Among them, the new parameter IND should be the product of the quadratic term coefficient in Formula (6) and the knowledge recombination variable coefficient in Formula (8). However, due to the different scales of the models represented by the two formulas, the Logit coefficient is the logarithmic odds ratio, and the linear regression coefficient is the unit change. The average marginal effect should be calculated to estimate the indirect effect of the change in the square unit of the independent variable on the logarithmic odds ratio of the dependent variable through the mediating variable. Therefore, this work uses the Bootstrap method to obtain the confidence interval of the parameter IND, as shown in Table 7. Since the distribution of the generated product is asymmetric, and the normal approximation depends on the symmetric distribution, the percentile interval or the deviation correction interval should be trusted because they do not depend on the symmetry assumption. In particular, the bias correction interval (BC) takes into account the estimated bias and is more reliable. From Table 7, the proportion of indirect effects is 0.00335, and the bias-corrected and percentile confidence intervals do not contain 0, indicating that the proportion is statistically significant. This means that the mediating variable (knowledge recombination) plays a significant mediating role in the inverted U-shaped relationship between the independent variable (collaboration

scale) and the dependent variable (breakthrough technological innovation). The indirect effect value is 0.00182, and all confidence intervals do not contain 0; the indirect effect is significant, that is, the influence of collaboration scale on breakthrough technological innovation through knowledge restructuring is positive and statistically reliable. The direct effect value is 0.542; all confidence intervals do not contain 0, and the direct effect is also significant. It shows that the inverted U-shaped relationship between collaboration scale and breakthrough technological innovation has both direct mechanisms and indirect mechanisms through knowledge recombination; that is, collaboration scale not only directly affects breakthrough technological innovation but also indirectly affects breakthrough technological innovation through knowledge recombination.

Table 7. Bootstrap mediating effect test.

	Observed Coefficient	**Bias**	**Bootstrap Std. Err.**	**[95% Conf. Interval]**	
_*bs*_1	0.00334808	−0.0000988	0.00171477	−0.0000128	0.006709 (N)
				0.0007594	0.0073632 (P)
				0.0009696	0.0078327 (BC)
_*bs*_2	0.00182208	0.0000942	0.00089701	0.000064	0.0035802 (N)
				0.0004518	0.0039478 (P)
				0.0004581	0.0039507 (BC)
_*bs*_3	0.54241064	0.0235249	0.07898583	0.3876013	0.69722 (N)
				0.426281	0.7399049 (P)
				0.4006867	0.6913674 (BC)
_*bs*_4	0.54421769	0.0235577	0.0793413	0.3887116	0.6997238 (N)
				0.4264621	0.742473 (P)
				0.4013731	0.6944869 (BC)

Note: _*bs*_1 is *r*(proportion), _*bs*_2 is *r*(ind_effect), _*bs*_3 is *r*(direct_effect), _*bs*_4 is *r*(total_effect). N is the normal approximation method, P is the percentile method, and BC is the deviation correction method.

However, the indirect effect value of the Bootstrap method is relatively small. In this work, knowledge recombination plays a nonlinear mediating role in collaboration scale and breakthrough technological innovation. The intermediary path of knowledge recombination mainly leads to a relatively narrow "optimal collaboration scale range", which can be found in the inverted U-shaped mediation effect of knowledge recombination. When using the Bootstrap method to calculate the average indirect effect of the entire sample, this effect that is strong within a specific interval but weak outside the interval is "averaged" to a smaller value (0.001822). Therefore, the relatively low value of the Bootstrap method in testing the indirect effects of knowledge recombination is an inevitable result of the "nonlinear mediation" model. In other words, a smaller proportion of average indirect effects does not mean that the knowledge recombination mechanism is unimportant but rather accurately reflects its conditional nature.

It is precisely due to the limitations of the Bootstrap method that the two-stage test method is further used to test the mediating effect of knowledge recombination. We further examine the effects of knowledge restructuring ability on radical technological innovation so as to provide direct evidence for the mediating effect. In Formula (6), the adjusted R^2 value is 0.041, and knowledge recombination is used as the explanatory variable. The regression results show that the first-order coefficient of the collaboration scale is positive, the second-order coefficient is negative, and both are significant at the 5% significance level, indicating that there is also an "inverted U-shaped" relationship between the collaboration scale and knowledge recombination. The extreme value point is 36.392, which is located in the 95% confidence interval of Fieller [18.867, 73.088]. In Formula (8), the mediating variable, knowledge recombination, is introduced, and the pseudo-R^2 becomes 0.200. Com-

pared with the pseudo-R^2 value of Formula (6), the model has a higher interpretation of the data and a better fitting effect. It can be seen that when the collaboration scale increases at the initial stage, the ability of knowledge recombination also increases. However, after exceeding a certain threshold, the growth of the collaboration scale inhibits breakthrough technological innovation, which verifies that knowledge recombination is the intermediary variable of the collaboration scale affecting breakthrough technological innovation. Therefore, by combining the Bootstrap method and two-stage method, it can be found that Hypothesis 2 (H2) is verified.

4.3. Endogeneity Test

Considering that breakthrough innovation results may in turn attract more collaborators and there is a two-way causal relationship, the selection of variables is limited, and key variables are omitted, as well as endogenous problems such as measurement errors. This work draws on the research of Kanama and Nishikawa [49] and uses instrumental variables in the Logit model to avoid this problem. The mean value of the collaboration scale of patents in the same year in each country or region was used as an instrumental variable of the collaboration scale, and the control function method was used for endogenous analysis. From the perspective of relevance, patents in the same year or region face a similar external environment, so they have certain relevance. The instrumental variable is a macro-level, structural feature that reflects a country's overall innovation cooperation in a specific year. In addition to the influence of the collaboration scale, the breakthrough of a single patent depends mainly on its own unique and micro-level factors. Therefore, the average collaboration scale at the national level will not directly determine whether a patent is a breakthrough patent, so it satisfies the principle of exogenousness. The choice of this instrumental variable is reasonable.

Firstly, the weak instrumental variable test was performed on the instrumental variables, where the F statistic = 390.43, which was much larger than the Stock-Yogo test 10% level critical value of 16.38, indicating that there was no weak instrumental variable problem. Then, a two-stage test is performed, and the results are shown in Table 8. In the first stage, the endogenous variables were regressed with the instrumental variables to generate the residual term v1, and the v1 coefficient was significant ($p < 0.05$), rejecting the original hypothesis and being endogenous. In the second stage, the residual term was added to the original Logit model to test its significance. The residual term v1 is significant at the 5% level, indicating that there was an endogenous problem. After using the control function method, the primary term coefficient increased from 0.065 to 0.159, and the occurrence ratio increased by exp (0.1592) $\approx$ 1.173 times for every 1 unit increase in the collaboration scale. For model fitting, the log-likelihood value of the control function method is -6102.11, and the log-likelihood value of the standard Logit regression is -6103.93. The difference between the log-likelihood values of the two methods is 1.8, which supports the control function method. The AIC values of the two models are further calculated. The AIC value of the control function method is 12,224.22, and the AIC value of the standard Logit regression is 12,225.87, indicating that the control function method is superior to the standard Logit method. Therefore, after endogenous control, the nonlinear effect of the core explanatory variables is still significant: the first-order coefficient of the collaboration scale is positive and significant at the 1% level, and the second-order coefficient is negative at the same significance level, which verifies that its "inverted U-shaped" relationship is still robust.

Table 8. Regression Results.

Variable	(1) Control Function Method		(2) Standard Logit Regression
	Phase 1 (CS)	Phase 2 (BTI)	BTI
CS		0.1592 ***	0.0651 ***
		(0.0487)	(0.0093)
CS^2		−0.0006 ***	−0.0006 ***
		(0.0002)	(0.0002)
IV	0.60678 ***		
	(0.00000)		
Cons_	4.51164 ***	−15.3327 ***	−14.6298 ***
	(0.00000)	(1.2167)	(1.1520)
Controls	YES	YES	YES
Residual Erro	−0.0960 **		
	(0.0489)		
N	35,955	35,955	35,955
Logarithmic Likelihood		−6102.11	−6103.93
$adj.R^2$	0.0505		
Pseudo R^2		0.1991	0.1989

The standard error is in parentheses. **, *** are significant at the level of 5% and 1%, respectively.

4.4. Robustness Test

To further verify the credibility and robustness of the empirical results, the substitution regression method and the control variable method are employed to assess robustness. First, the substitution regression method is used. In this work, the Probit regression method is employed to replace the Logit regression analysis method, thereby avoiding the deviation in results caused by the regression method. The results in Table 9a show that, with an increase in the collaboration scale, the level of breakthrough technological innovation first increases and then decreases. The results of Probit regression are highly consistent with the previous conclusions, and there is no significant difference, which confirms that the empirical results of this work have good robustness. Secondly, the method of increasing control variables is used. This work adds the variable of transnational collaboration (TC), and the empirical results are shown in Table 9b. The first term of the collaboration scale is positive and the second term is negative, both of which are significant at the level of 5%, which proves that the relationship between the collaboration scale and breakthrough technological innovation is still inverted U-shaped, which further verifies the robustness of the research results.

Table 9. Robustness Tests.

(a) Replacement Regression Method		(b) Increased Control Variables	
	BTI		BTI
CS	0.03173 ***	CS	0.06585 ***
	(0.00000)		(0.00000)
CS^2	−0.00027 **	CS^2	−0.00056 *
	(0.01214)		(0.02044)
Cons_	−6.62382 ***	Cons_	−14.61440 ***
	(0.00000)		(0.00000)
Controls	YES	Controls	YES
N	35,955	N	35,955
Pseudo R^2	0.202	Pseudo R^2	0.199

The robust standard error is in parentheses. *, **, *** are significant at the 10%, 5%, 1% level, respectively.

4.5. Heterogeneity Analysis

To deeply understand the differences in the influence of collaboration scale in different countries and regions with different economic strengths, this work counted the five countries with the highest number of technology patents in the field of recommendation systems, including China, the United States, Japan, South Korea, and Russia, as the research objects. The heterogeneity analysis grouped patents based on the nationality of the first applicant (for example, patents with a Chinese entity as the first applicant were assigned to the "China" group) to compare differences among core innovation entities from different countries. In this paper, according to the above grouping, the patents of each group are counted. The national patent data of the top 5 countries in the number of patents are selected for regression, and finally, the heterogeneity analysis is carried out. The heterogeneity test results of each country are shown in Table 10.

Table 10. Results of heterogeneity analysis.

Variable	China	The United States	Japan	South Korea	Russia
			BTI		
CS	0.19628 ***	0.08013 ***	0.04333	0.12849 **	2.06535 ***
	(0.00000)	(0.00000)	(0.30967)	(0.01817)	(0.00011)
CS^2	−0.00698 ***	−0.00070 **	−0.00017	−0.00077	−0.21492 ***
	(0.00069)	(0.01742)	(0.93789)	(0.48078)	(0.00108)
Cons_	−28.54177	−98.11454 ***	−171.69288 ***	51.22654 **	13.18525
	(0.11412)	(0.00000)	(0.00000)	(0.03400)	(0.42950)
Controls	YES	YES	YES	YES	YES
N	18,617	6074	3370	2970	2265
Pseudo R^2	0.174	0.025	0.091	0.160	0.283

The robust standard error is in parentheses. **, *** are significant at the 5%, 1% level, respectively.

The data in Table 10 reveal the differences in the relationship between collaboration scale and breakthrough technological innovation in five countries, including China, the United States, Japan, the United Kingdom, and Russia. It is found that the collaboration scale of China, the United States, and Russia has a significant inverted U-shaped relationship with breakthrough technological innovation, but the critical points of their optimal collaboration scale are different: the optimal thresholds are 14.058 entities for China, 57.151 entities for the United States, and only 4.801 entities for Russia. In contrast, the data of Japan and South Korea show that the influence of collaboration scale and its square term on breakthrough technological innovation is not significant, indicating that there is no clear linear or nonlinear relationship.

The higher critical point of the United States benefits from its mature industry-university-research innovation ecology and strong organizational coordination ability. The U.S. government supports close collaboration between schools and enterprises, cultivates and encourages entrepreneurship and innovation, the transformation of scientific research projects, and forms an internal and external technology transformation service system and an industry-university-research integration ecological science and technology system, which can maintain efficiency in a larger collaboration network to coordinate a higher optimal collaboration scale [50]. The critical point of China is significantly lower than that of the United States, which is mainly attributed to the fact that China's innovation system is still in the development stage [51], the organizational coordination ability and efficiency of the collaboration network are limited, so that large-scale cooperation is easy to cause "free rider" behavior or management confusion [52], so the optimal collaboration scale is relatively small. The low critical point of Russia is because Russian enterprises have

low enthusiasm for technological innovation and are more inclined to purchase mature foreign technical solutions, so their cooperation mechanism is not flexible enough [53].

The collaboration scale has no significant influence on Japan and South Korea's breakthrough technology research. The possible reason is that the breakthrough technological innovations in Japan and South Korea are mostly independent research and development (R&D) rather than cooperation. Among them, only 30 out of 283 breakthrough technologies in Japan are cooperative R&D (accounting for only 10.6%), and only 8 out of 30 breakthrough technologies in South Korea are cooperative R&D (including 5 non-organizational cooperations, accounting for only 13.33%). Furthermore, both have shown a highly engineering-oriented collaboration model (only inter-enterprise cooperation), rather than industry-university-research cooperation. Only 3 out of 30 cooperative R&D projects in Japan are industry-university-research cooperation, accounting for only 10%, and only 1 out of 8 cooperative R&D projects in South Korea is industry-university-research cooperation, accounting for only 12.5%. From the above two aspects, it can be found that Japan and South Korea's breakthroughs in the field of recommendation systems rely more on independent enterprise technology development, and even in cooperative R&D, they rely more on joint R&D between enterprises rather than industry-university-research cooperation, which inevitably limits the emergence of cross-border knowledge recombination. In previous studies, some scholars have also pointed out that R&D funds in Japan or South Korea are mostly invested in leading enterprises [54,55], which leads to certain closed innovation characteristics of enterprises. Meanwhile, studies have shown that both Japan and South Korea are greatly influenced by domestic conglomerates in their technology research and development [56]. The internal cooperation between the two countries' conglomerates is relatively closed, and companies have formed a close community of interests through cross shareholding and personnel dispatch [57,58], even resorting to "nepotism" under the requirements of the chaebol system [55]. Although this model facilitates resource sharing, it also leads to a lower acceptance of external innovative technologies and concepts by enterprises. Therefore, the limited cooperation between Japan and South Korea has led to a phenomenon where the collaboration scale has no significant influence on breakthrough technological innovation.

5. Conclusions and Future Work

5.1. Research Conclusions

In the context of the surge in the cost of cutting-edge technology breakthroughs and the extension of the transformation cycle, the world's major economies have formed an innovative ecology of multi-agent collaboration and all-factor integration by building a multi-level international innovation alliance [37]. This cross-regional and cross-domain collaborative innovation network not only provides a new paradigm for overcoming breakthrough technologies but also injects sustainable momentum into the transformation and upgrading of technological systems [40]. As the key to technological research, human resources and intellectual security brought by collaborative innovation are important ways to optimize resource allocation [21]. Whether breakthrough technological innovation follows the law of "more people, more power" still needs to be explored.

Given this, this work focuses on the stage of breakthrough technological innovation from "0 $\rightarrow$ 1", and explores the influence mechanism of collaboration scale on breakthrough technological innovation. To ensure the scientificity and standardization of the research, the technology of 58 countries and regions in the field of recommendation systems is taken as the research object, and the related technology patents are taken as the main data carrier to test the direct effect of collaboration scale on breakthrough technological innovation and the mediating effect of knowledge recombination. This work proposes a relationship

model between collaboration scale and breakthrough technological innovation based on patent data from the recommendation system field. It aims to elucidate how collaboration scale influences breakthrough technological innovation through knowledge restructuring, thereby providing theoretical support and practical guidance for enterprises, institutions, and governments in innovation activities to advance technological innovation. The core finding of this work is that a significant inverted U-shaped relationship exists between collaboration scale and breakthrough technological innovation. This conclusion aligns with some prior research perspectives while revealing deeper distinctions and unique contributions of this work. On the one hand, it supports the findings of An and Shan [59] and Deng et al. [47]. An and Shan [59] used kernel density estimation to find that output across disciplines exhibits a "Kuznets curve" pattern with team size, while Deng et al. [47] also noted medium-sized teams offer advantages in balancing efficiency and diversity. This work differs by deepening the measurement of core variables and research perspective. Previous studies predominantly used the number of inventors as a single metric for collaboration scale. In contrast, this work innovatively constructs a dual-dimensional metric combining the number of inventors and the number of institutions. This dual approach not only more comprehensively captures the complexity of collaborative networks but also effectively avoids the potential bias of a single metric. Furthermore, unlike studies such as An and Shan [59] that conduct broad multidisciplinary scans, this research focuses on the cutting-edge and highly technologically representative field of "recommendation systems". By employing a larger patent sample for in-depth analysis, the findings gain greater specificity and practical relevance within this particular technological context.

Second, this work reveals that "knowledge recombination" mediates the relationship between collaboration scale and breakthrough technological innovation. This finding resonates with the conclusion from Zhong et al. [11] that knowledge recombination creates conditions conducive to breakthrough innovation. However, existing studies predominantly analyze knowledge recombination as an independent variable directly influencing innovation [11] or focus on the roles of different knowledge sources across distinct stages of the technology lifecycle [60]. Few studies systematically examine the nonlinear influence of collaboration scale on the efficiency of knowledge recombination and the ultimate innovation output.

This work leverages patent data from 35,955 patents across 58 countries and regions within the recommendation system technology domain. Its substantial sample size and cross-border characteristics enhance the statistical validity and generalizability of the findings. In contrast, some prior research has focused on specific industries within a single country [11] or two technology domains within a specific country [60]. Finally, this work employs a stricter definition of "breakthrough technological innovation". By dynamically screening the top 5% of patents by citation volume annually, it ensures the research subjects represent genuinely high-influence technological leaps. More importantly, the heterogeneity analysis reveals significant variations in optimal collaboration scale across nations, providing crucial empirical evidence for countries to develop differentiated policies on technological innovation and talent cooperation—a novel perspective previously unexplored in the literature.

5.2. Management Implications

(1) Optimize the collaboration scale and formulate regional differentiated governance strategies. According to the research conclusions of this paper, the scale of cooperation and breakthrough technological innovation shows an inverted U-shaped relationship, which indicates that there is an optimal range for the scale of R&D cooperation. Based on this, managers should optimize the scale of cooperation and avoid blindly

pursuing the number of partners. Instead, they should evaluate the scale of existing cooperation and regularly audit the number of cooperation projects and partners. If the project is excessive, it may lead to management distraction and resource dilution, which is not conducive to breakthrough innovation output. In breakthrough innovation, R&D management departments need to allocate resources reasonably in internal and external innovation activities, establish a dynamic evaluation mechanism, use the inverted U-shaped relationship between collaboration scale and breakthrough technological innovation, evaluate the marginal changes in cooperation income and coordination cost loss in real time, and dynamically adjust the cooperation structure according to the innovation stage. According to the results of heterogeneity analysis, we propose differentiated policies and establish different collaboration scale optimization mechanisms in different countries and regions. Chinese enterprises should focus on establishing deep and mutual trust strategic alliances with a few core partners, rather than spreading the net. American companies can build and lead large innovation alliances more confidently and take advantage of their ecological advantages. Russia should encourage enterprises to try to establish small-scale, pilot cooperation projects with external institutions (especially universities). Japanese and South Korean enterprises should deeply analyze the characteristics of their own innovation ecology and find other core variables that affect their breakthrough technological innovation. Countries need to focus on supporting the cutting-edge exploration of elite teams, optimizing the environment for team operation, creating an incentive compatibility mechanism for knowledge flow, and promoting knowledge flow efficiency to promote breakthrough technological innovation. At the same time, while encouraging cooperation, it is necessary to focus on improving the basic conditions of cooperation, such as infrastructure, institutional environment, and capacity building, to reduce the "threshold" of effective cooperation and avoid simply pursuing scale at the expense of efficiency.

(2) Build a knowledge recombination driving mechanism to promote the high-quality development of breakthrough technological innovation. It is necessary to shift from simply encouraging "expanding collaboration scale" to "optimizing the cooperation structure and enabling the recombination process", so that the knowledge heterogeneity contained in the large-scale cooperation network can be effectively identified, collated, integrated, and transformed into breakthrough technological innovation results. In order to promote the construction of the driving mechanism of knowledge restructuring, countries and governments need to support the construction of large-scale cooperative knowledge infrastructure investment to reduce the transaction costs of restructuring; encourage open data platforms; and promote the conditional opening of non-classified data of governments, universities, and enterprises on the platform under the premise of ensuring data security and privacy so as to provide assistance for cross-border research. Establish an expert database, use AI technology to intelligently recommend partners with complementary knowledge according to project requirements, solve the problem of "how to find the right collaborators", and achieve talent matching. In addition, the organization is the direct place of knowledge recombination. Therefore, the organization needs to break the internal door wall and stimulate the recombination behavior. Encourage researchers to work in laboratories or enterprises in other disciplines for short-term work, experience different knowledge systems first-hand, and stimulate restructuring inspiration; researchers with different knowledge backgrounds are encouraged to communicate and cooperate with each other to create innovative sparks. Encourage enterprises to build R&D entities with deep integration with universities or research institutes.

Enterprise engineers can stay in university laboratories. University researchers regularly visit enterprise R&D centers and production lines to achieve deep restructuring of "scientific knowledge" and "technical knowledge". Under the open innovation paradigm, it is necessary to promote the construction of knowledge recombination ability from three dimensions. Firstly, a new technology evolution path is constructed, and a cross-domain knowledge interaction platform is established to promote the deconstruction of the knowledge system and achieve reintegration. Secondly, focus on building a competitive advantage in the technology ecosystem. Through multiple strategies such as patent layout, standard formulation, and demonstration application, the accumulation and release of technological potential energy are accelerated, and the technological dominance is established to provide inheritable knowledge assets and technological paradigms for subsequent iterative innovation so as to effectively reduce the cost of innovation trial and error. Finally, while exploring new technology areas, it is also necessary to focus on the strategic management of existing knowledge assets, such as constructing a dynamic knowledge management capability system, including establishing a cross-time domain knowledge traceability mechanism, a cross-domain knowledge association network, and an intelligent knowledge matching system; realizing the analysis and identification of massive patent documents and technical documents; finding potential portfolio innovation opportunities; transforming discrete technical modules into new technical solutions with market value; and promoting the high-quality development of breakthrough technological innovation.

(3) Promote transnational technical cooperation, share opportunities, and promote development. Due to the differences in systems and development among countries, the inverted U-shaped effect of collaboration scale on breakthrough technological innovation is different in different countries. In order to promote global breakthrough technological innovation, countries and regions need to work closely together to establish transnational collaboration centers and patent sharing pools in key areas to promote new technologies. Deeply coupling the advantages of knowledge recombination with the advantages of application scenarios, balancing the interests of all countries, and ensuring the sustainability of collaboration. At the same time, in order to break through the existing institutional bottleneck of transnational cooperation, it is necessary to establish and improve various guarantee mechanisms, strengthen the formulation and implementation of breakthrough technological innovation strategies, share risks, share opportunities, and ultimately achieve common development.

5.3. Limitations

Although this work provides some valuable insights, there are still some limitations. First of all, this work only uses the most widely used measure of patent citations as a measure of breakthrough technological innovation. Although easy to obtain and quantify, it cannot cover all forms of innovation. Secondly, the sample is limited to the patent data in the recommendation system field, which may limit the universality of the research conclusions in other fields. Finally, this work only explores the process of breakthrough technological innovation from 0 to 1, which lacks the exploration of the whole cycle of breakthrough technological innovation from generation to promotion and application.

5.4. Prospect of Working

In the next work plan, the research contents of three aspects are to be carried out. First of all, we will focus on other technical fields of artificial intelligence, such as natural language processing, computer vision, and expert systems, and continue to explore the

internal mechanism of breakthrough technological innovation. Secondly, it focuses on the promotion process of breakthrough technological innovation from "1 → *N*", and compares and analyzes the process of "0 → 1" explored in this work from scratch so as to realize the complete innovation mechanism and path of breakthrough technology in the technology system. At the same time, this work explores the heterogeneity of the collaboration scale on breakthrough technological innovation in different countries, and we will reveal the causes of heterogeneity in future work.

Author Contributions: Conceptualization, L.S.; methodology, L.S. and S.C.; software, S.C.; validation, L.S., X.Y. and J.L.; formal analysis, L.S. and S.C.; investigation, L.S.; resources, L.S. and X.Y.; data curation, L.S.; writing—original draft preparation, S.C., L.S. and J.L.; writing—review and editing, L.S. and X.Y.; visualization, S.C.; supervision, L.S.; project administration, L.S.; funding acquisition, L.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China [72304101, 72171136], the Guangdong Basic and Applied Basic Research Foundation [2024A1515011559], and the Guangdong Provincial Philosophy and Social Science Planning Project [GD24YGL22].

Data Availability Statement: Data derived from public domain resources. These data were derived from the following resources available in the public domain: https://www.patentguru.com/cn; https://www.incopat.com/; https://data.worldbank.org.cn/.

Acknowledgments: We would like to express our gratitude to Yinghong Ma from Shandong Normal University for providing revision suggestions for this research. At the same time, we also want to thank the editors for their careful handling of this manuscript. Last but not least, we would like to thank the anonymous reviewers for carefully reviewing this manuscript and providing valuable revision suggestions, greatly improving the scientific accuracy and readability of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Sharif, M.N. Technological innovation governance for winning the future. *Technol. Forecast. Soc. Chang.* **2012**, *79*, 595–604. [CrossRef]
2. Winnink, J.J.; Tijssen, R.J.; Van Raan, A.F.J. Searching for new breakthroughs in science: How effective are computerised detection algorithms? *Technol. Forecast. Soc. Chang.* **2019**, *146*, 673–686. [CrossRef]
3. Burgelman, R.A.; Siegel, R.E. Cutting the strategy diamond in high-technology ventures. *Calif. Manag. Rev.* **2008**, *50*, 140–167. [CrossRef]
4. Datta, A.A.; Srivastava, S. (Re)conceptualizing technological breakthrough innovation: A systematic review of the literature and proposed framewor. *Technol. Forecast. Soc. Chang.* **2023**, *194*, 122740. [CrossRef]
5. McDermott, C.M.; O'Connor, G.C. Managing radical innovation: An overview of emergent strategy issues. *J. Prod. Innov. Manag.* **2002**, *19*, 424–438. [CrossRef]
6. Ahuja, G.; Katila, R. Technological acquisitions and the innovation performance of acquiring firms: A longitudinal study. *Strateg. Manag. J.* **2001**, *22*, 197–220. [CrossRef]
7. Grigoriou, K.; Rothaermel, F.T. Structural microfoundations of innovation: The role of relational stars. *J. Manag.* **2001**, *40*, 586–615. [CrossRef]
8. Yin, X.; Song, L.; Pan, Y. Diversity in knowledge search paths and breakthrough technological innovation: Evidence from new energy vehicle firms. *PLoS ONE* **2025**, *20*, e0326736. [CrossRef] [PubMed]
9. Xie, X.; Wang, L.; Zeng, S. Inter-organizational knowledge acquisition and firms' radical innovation: A moderated mediation analysis. *J. Bus. Res.* **2018**, *90*, 295–306. [CrossRef]
10. Kamuriwo, D.S.; Baden-Fuller, C.; Zhang, J. Knowledge development approaches and breakthrough innovations in technology-based new firms. *J. Prod. Innov. Manag.* **2017**, *34*, 492–508. [CrossRef]
11. Zhong, C.; Huang, R.; Duan, Y.; Sunguo, T.; Dello Strologo, A. Exploring the impacts of knowledge recombination on firms' breakthrough innovation: The moderating effect of environmental dynamism. *J. Knowl. Manag.* **2024**, *28*, 698–723. [CrossRef]

12. Li, C.; Zhang, X.; Wang, L. How Industry–University–Research Integration Promotes Green Technology Innovation in Chinese Enterprises: The Dual Mediating Pathways and Nonlinear Effects. *Systems* **2025**, *13*, 696. [CrossRef]
13. Xu, L.; Li, J.; Zhou, X. Exploring new knowledge through research collaboration: The moderation of the global and local cohesion of knowledge networks. *J. Technol. Transf.* **2019**, *44*, 822–849. [CrossRef]
14. Thomas, R.J.; Wind, Y. Symbiotic innovation: Getting the most out of collaboration. In *Evolution of Innovation Management: Trends in an International Context*; Brem, A., Viardot, É., Eds.; Palgrave Macmillan: London, UK, 2013; pp. 1–31.
15. Schumpeter, J.A. *Business Cycles*; McGraw-Hill: New York, NY, USA, 1939; pp. 411–413.
16. Fleming, L. Recombinant uncertainty in technological search. *Manag. Sci.* **2001**, *47*, 117–132. [CrossRef]
17. Nelson, R.R.; Winter, S.G. An Evolutionary Theory of Economic Change. In *Resources, Firms, and Strategies: A Reader in the Resource-Based Perspective*; Foss, N.J., Ed.; Oxford University Press: Oxford, UK, 1997.
18. Lucas, R.E. Understanding Business Cycles. In *Essential Readings in Economics*; Estrin, S., Marin, A., Eds.; Macmillan Education: London , UK, 1995; pp. 306–327.
19. Ahuja, G.; Morris, L.C. Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strateg. Manag. J.* **2001**, *22*, 521–543. [CrossRef]
20. Wu, K.; Sun, J.; Wang, J.; Kang, L. How does science convergence influence technology convergence? Different impacts of science-push and technology-pull. *Technol. Forecast. Soc. Chang.* **2015**, *215*, 124114. [CrossRef]
21. Karim, S.; Kaul, A. Structural recombination and innovation: Unlocking intraorganizational knowledge synergy through structural change. *Organ. Sci.* **2015**, *26*, 439–455. [CrossRef]
22. Song, L.; Zhu, G.; Yin, X. Evaluating the wisdom of scholar crowds from the perspective of knowledge diffusion. *Scientometrics* **2024**, *129*, 5103–5139. [CrossRef]
23. Newman, M.E.J. Scientific collaboration networks. I. Network construction and fundamental results. *Strateg. Manag. J.* **2001**, *64*, 016131. [CrossRef]
24. Wu, L.; Wang, D.; Evans, J.A. Large teams develop and small teams disrupt science and technology. *Nature* **2019**, *566*, 378–382. [CrossRef]
25. Fang, S.; Qin, Y. Social Identification in Open Innovation Projects: Role of Knowledge Collaboration and Resource Interdependence. *Systems* **2025**, *13*, 129. [CrossRef]
26. Coleman, J.S. Social Capital in the Creation of Human Capital. *Am. J. Sociol.* **1988**, *94*, S95–S120. [CrossRef]
27. Grover, V.; Malhotra, M.K. Transaction cost framework in operations and supply chain management research: Theory and measurement. *J. Oper. Manag.* **2003**, *21*, 457–473. [CrossRef]
28. Simon, H.A. *Administrative Behavior: A Study of Decision Making Processes in Administrative Organization*, 4th ed.; Free Press: New York, NY, USA, 1997; p. 124.
29. Soda, G.; Mannucci, P.V.; Burt, R.S. Networks, creativity, and time: Staying creative through brokerage and network rejuvenation. *Acad. Manag. J.* **2021**, *64*, 1164–1190. [CrossRef]
30. Teixeira, S.J.; Veiga, P.M.; Fernandes, C.A. The knowledge transfer and cooperation between universities and enterprises. *Knowl. Manag. Res. Pract.* **2019**, *17*, 449–460. [CrossRef]
31. Liu, X.; Bollen, J.; Nelson, M.L.; Van de Sompel, H. Co-authorship networks in the digital library research community. *Inf. Process.* **2005**, *41*, 1462–1480. [CrossRef]
32. Gruber, M.; Harhoff, D.; Hoisl, K. Knowledge Recombination Across Technological Boundaries: Scientists vs. Engineers. *Manag. Sci.* **2013**, *59*, 837–851. [CrossRef]
33. Yang, N.; Song, Y.; Zhang, Y.; Wang, J. Dark side of joint R&D collaborations: Dependence asymmetry and opportunism. *J. Bus. Ind. Mark.* **2020**, *35*, 741–755. [CrossRef]
34. Kumar, P.; Zaheer, A. Ego-network stability and innovation in alliances. *Acad. Manag. J.* **2019**, *62*, 691–716. [CrossRef]
35. Boh, W.F.; Ren, Y.; Kiesler, S.; Bussjaeger, R. Expertise and collaboration in the geographically dispersed organization. *Organ. Sci.* **2007**, *18*, 595–612. [CrossRef]
36. Arts, S.; Veugelers, R. Technology familiarity, recombinant novelty, and breakthrough invention. *Ind. Corp. Chang.* **2015**, *24*, 1215–1246. [CrossRef]
37. Wang, T.; Wang, J.; Shi, J.; Sun, J.; Kang, L. Technological recombinant strategy and breakthrough innovation of team: The moderating role of science linkage. *J. Inf.* **2025**, *19*, 101613. [CrossRef]
38. Vakili, K.; Kaplan, S. Organizing for innovation: A contingency view on innovative team configuration. *Strateg. Manag. J.* **2021**, *42*, 1159–1183. [CrossRef]
39. Fontana, R.; Nuvolari, A.; Shimizu, H.; Vezzulli, A. Reassessing patent propensity: Evidence from a dataset of R&D awards, 1977–2004. *Econ. Innov. Hist. Perspect. Honour Nick Von Tunzelmann* **2013**, *42*, 1780–1792.
40. Gao, Y.; Hu, Y.; Liu, X.; Zhang, H. Can public R&D subsidy facilitate firms' exploratory innovation? The heterogeneous effects between central and local subsidy programs. *Res. Policy* **2021**, *50*, 104221. [CrossRef]

41. Erbey, A.; Gündüz, C.; Fidan, Ü. Digitalization, Sustainability, and Radical Innovation: A Knowledge-Based Approach. *Sustainability* **2025**, *17*, 2972. [CrossRef]
42. Olechowski, A.L.; Eppinger, S.D.; Joglekar, N.; Tomaschek, K. Technology readiness levels: Shortcomings and improvement opportunities. *Syst. Eng.* **2020**, *23*, 395–408. [CrossRef]
43. Alvarado, R.; Murshed, M.; Cifuentes-Faura, J.; Işık, C.; Hossain, M.R.; Tillaguango, B. Nexuses between rent of natural resources, economic complexity, and technological innovation: The roles of GDP, human capital and civil liberties. *Resour. Policy* **2023**, *85*, 103637. [CrossRef]
44. Audretsch, D.B.; Belitski, M.; Chowdhury, F. Knowledge investment and search for innovation: Evidence from the UK firms. *J. Technol. Transf.* **2024**, *49*, 1387–1410. [CrossRef]
45. Jiang, Y.Y.; Liu, X.L.; Wang, L.Y. The Disruptive Innovation Evaluation and Empirical Analysis of Chinese, Japanese, Indian, and South Korean Scientific Journals. *J. Sch. Publ.* **2025**, *56*, 60–78. [CrossRef]
46. Svensson, R. Patent value indicators and technological innovation. *Empir. Econ.* **2022**, *62*, 1715–1742. [CrossRef]
47. Deng, C.; Zhang, X.; Yang, Z.; Zhang, M. Impact of interdisciplinarity on disruptive innovation: The moderating role of collaboration pattern and collaboration size. *Scientometrics* **2025**, *130*, 2379–2401. [CrossRef]
48. Lind, J.T.; Mehlum, H. With or without U? The appropriate test for a U-shaped relationship. *Oxf. Bull. Econ. Stat.* **2010**, *72*, 109–118. [CrossRef]
49. Kanama, D.; Nishikawa, K. What type of obstacles in innovation activities make firms access university knowledge? An empirical study of the use of university knowledge on innovation outcomes. *J. Technol. Transf.* **2017**, *42*, 141–157. [CrossRef]
50. Gray, D.O. Cross-sector research collaboration in the USA: A national innovation system perspective. *Sci. Public Policy* **2011**, *38*, 123–133. [CrossRef]
51. Singh, A.; Wong, P.K.; Ho, Y.P. The role of universities in the national innovation systems of China and the East Asian NIEs: An exploratory analysis of publications and patenting data. *Asian J. Technol. Innov.* **2015**, *23*, 140–156. [CrossRef]
52. Song, L.; Ma, Y. Evaluating tacit knowledge diffusion with algebra matrix algorithm based social networks. *Appl. Math. Comput.* **2022**, *428*, 127125. [CrossRef]
53. Fonotov, A. The role of communications in the development of a national innovation system. *Econ. Math. Methods* **2020**, *56*, 35–44. [CrossRef]
54. Lynn, L.H. Recovering from success: Innovation and technology management in Japan. *J. Jpn. Stud.* **2008**, *34*, 543–548. [CrossRef]
55. Cooke P. A ground-up "Quaternary" innovation strategy for South Korea using entrepreneurial ecosystem platforms. *J. Open Innov. Technol. Mark. Complex.* **2017**, *3*, 1–16. [CrossRef]
56. Sakakibara, M.; Cho, D.S. Cooperative R&D in Japan and Korea: A comparison of industrial policy. *Res. Policy* **2002**, *31*, 673–692. [CrossRef]
57. Munro, J.F. Publicly Funded Research and Innovation in Japan and the Outlook for International Cooperation. In *Publicly Funded Transport Research in the PR China, Japan, and Korea: Policies, Governance and Prospects for Cooperation with the Outside World*; Springer International Publishing: Cham, Switzerland, 2017; pp. 543–548.
58. Ahn, J.M.; Kim, D.; Moon, S. Determinants of innovation collaboration selection: A comparative analysis of Korea and Germany. *Innovation* **2017**, *19*, 125–145. [CrossRef]
59. An, R.; Shan, W. The exploration in the size of scientific collaboration team using kernel density estimation. *Aslib J. Inf. Manag.* **2023**, *75*, 821–843. [CrossRef]
60. Kalthaus, M. Knowledge recombination along the technology life cycle. *J. Evol. Econ.* **2020**, *30*, 643–704. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.