

Article

When Do Structural Holes Yield Breakthrough Innovation? An Inverted U-Shape Bounded by Collaboration-Layer Centralities [†]

Shugang Li ¹, Jinxian Dong ¹, Zhaoxu Yu ², Zhifang Wen ¹, Mengsi Sun ^{1,*} and Xinyi Ye ^{1,*}

¹ School of Management, Shanghai University, Shanghai 200444, China; luck_li@shu.edu.cn (S.L.); dongjinxian@shu.edu.cn (J.D.); 18112094238@163.com (Z.W.)

² Department of Automation, East China University of Science and Technology, Shanghai 200231, China; yyzx@ecust.edu.cn

* Correspondence: sunmengsi@shu.edu.cn (M.S.); ye_xinyi@shu.edu.cn (X.Y.)

[†] This paper is an extended version of our paper published in 5th International Communication Engineering and Cloud Computing Conference, CECCC 2023, Nanjing, China, 27–29 October 2023. S. Li, Z. Wen. Interaction Effect of Knowledge Network and Collaboration Network on Breakthrough Innovation.

Abstract

Breakthrough innovation—central to industrial competitiveness and the ongoing clean-energy transition—remains persistently constrained by information homogenization and weak cross-domain integration in single-layer innovation networks. Technology Innovation Composite Networks (TICNs) have therefore been advocated as dual-layer platforms coupling knowledge and collaboration networks, yet the cross-layer mechanism through which they generate breakthrough outputs has not been specified. This paper specifies and tests how knowledge-layer structural holes open access to heterogeneous information that must cross into the collaboration layer to be recombined into breakthroughs. Two distinct boundaries shape the outcome. Inventors' finite cognitive processing capacity makes integration returns decay along an inverted U-shape; separately, excessive degree and closeness centrality drive the collaboration layer into homogenization and localization, narrowing the range of structural holes it can productively absorb and shifting the breakthrough peak toward lower structural-hole levels. Together, they delineate an optimal cross-layer integration zone. Using panel data on 10,681 patents, 948 inventors, and 5631 inventor-year observations from new energy (2004–2018), a fixed-effects negative binomial model confirms the inverted U-shape and the steepening, peak-shifting moderations of degree and closeness centrality; a Lind–Mehlum test places the turning point inside the observed data range, and negative binomial (robust SE), Poisson and zero-inflated Poisson specifications—together with a stricter top-1% breakthrough threshold—yield consistent results. The study moves multilayer network research from structural description toward mechanism-level identification and offers actionable network-design guidance.

Academic Editor: Luna Leoni

Received: 06 May 2026

Revised: 19 June 2026

Accepted: 24 June 2026

Published: 27 June 2026

Keywords: breakthrough innovation; Technology Innovation Composite Network; cross-layer integration mechanism; knowledge-network structural holes; collaboration-network centrality

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Breakthrough technological innovation has long been recognized as a key engine of industrial upgrading and national competitiveness, and prior network-based research has

begun to examine how the interaction between knowledge networks and collaboration networks shapes breakthrough innovation [1]. Countries around the world have accordingly invested heavily in building the network infrastructure that supports technological innovation. In recent years, the United States, the European Union and China have actively built cross-border knowledge exchange platforms: the U.S. National AI R&D Strategic Plan, China's Artificial Intelligence Development Plan, together with the Belt and Road Initiative Science and Technology Cooperation Platform, and the EU's Horizon 2020 program have all encouraged synergistic, transnational technological cooperation [2,3]. Riding on these platforms, global patent filings and cross-border co-inventions have grown rapidly over the past two decades, placing technological innovation at the center of national strategy.

Yet, despite this quantitative expansion, breakthrough technological innovation—the type of innovation that reshapes industries and the socio-economic structure [4,5], exemplified by high-efficiency solar power and advanced wind turbine technology—remains persistently constrained. The fundamental reason is that, within single-layer knowledge networks, information exchange is largely confined to a single technological domain or an existing technological framework, and inventors lack effective channels for cross-domain integration. Information homogenization and weak cross-domain integration, therefore, jointly depress the upper bound of what knowledge-network-based innovation can deliver: more patents are filed, but the share that genuinely reshapes a technological trajectory does not rise in step. This is the long-standing bottleneck against which this paper is directed.

To alleviate this bottleneck, Technological Innovation Composite Networks (TICNs) have emerged as a platform that couples the knowledge network and the collaboration network into a dual-layer system [6]. TICNs are conceptually attractive precisely because heterogeneous knowledge identified in the knowledge layer must be operationalized by inventors working together in the collaboration layer; without such cross-layer coupling, heterogeneous information remains latent and cannot be translated into breakthrough outputs. However, although TICNs are widely believed to promote breakthrough innovation, the specific cross-layer mechanism through which they do so remains poorly understood. Existing work is concentrated on single-layer analysis—typically either a knowledge network or a collaboration network in isolation—and offers limited insight into how structural features of one layer are conditioned by, and must pass through, the other [7]. The dual-layer architecture has been described, but the pathway by which information actually crosses layers has not been opened.

This paper opens that pathway by taking knowledge-network structural holes as the entry point and, on the TICN platform, providing a structural microfoundation for how heterogeneous information is converted into breakthrough innovation through cross-layer integration. Structural hole theory (SHT) holds that nodes spanning structural holes gain privileged access to novel, non-redundant information from otherwise disconnected knowledge sources [7]. In the TICN framework, however, structural holes no longer function merely as within-layer access points: they open the entry channels through which inventors acquire heterogeneous information, and that information must then flow across layers into the collaboration network before it can be recombined, negotiated and materialized into breakthrough outputs. This cross-layer pathway cannot be captured by single-layer analyses, which is precisely why Multilayer Networks Theory [8,9] is needed as the analytical scaffold: it is the only framework in which the joint operation of two layers—and the conditions under which information transfer between them succeeds or fails—can be formally represented.

The cross-layer integration mechanism, however, is bounded by two distinct forces that jointly shape a non-monotonic outcome. First, inventors have finite cognitive

capacity. Within a moderate range, additional structural holes expand the variety of heterogeneous inputs and raise breakthrough potential; once the number of structural holes surpasses an inventor's cognitive processing threshold, information overload and redundancy are triggered, and the integration return decays along an inverted U-shaped curve. Second, whether heterogeneous information can effectively cross layers depends on the state of the receiving layer itself. Two typical but opposing scenarios illustrate this point:

- Scenario 1: In highly centralized collaboration networks, inventors enjoy fast access to information, but innovation paths tend to converge. In biotechnology R&D, for example, high-centrality firms concentrate on a few key technology areas, producing bottlenecks in technology development [10].
- Scenario 2: In collaboration networks characterized by high closeness centrality among neighboring nodes, information flows quickly within local cliques but rarely reaches distant domains. Cross-industry SMEs, by contrast, often produce more groundbreaking and diverse innovations despite slower information flow [11].

These two scenarios reveal that excessively high degree centrality and closeness centrality in the collaboration layer drive the collaboration network itself into information homogenization and localization, stripping it of the capacity to receive heterogeneous information from the knowledge layer. When the receiving layer is degraded in this way, the cross-layer integration pathway is narrowed, and the structural-hole level at which the inverted U-shaped main effect peaks is pushed lower and the productive window narrows—on top of the decay that is already implied by the cognitive threshold. Taken together, the decaying main effect of structural holes and the dual-gate moderating effects of collaboration-layer centralities precisely delineate an optimal cross-layer integration zone, within which inventors reach the maximum level of breakthrough innovation.

Building on this logic, the present paper makes three connected contributions, each addressing a distinct gap left by single-layer analyses. Theoretically, it extends structural hole theory and multilayer network theory by formulating and testing a cross-layer integration mechanism that cannot be reduced to any single layer: structural holes are repositioned from a within-layer information advantage to a cross-layer entry channel, and collaboration-layer centralities are repositioned from access indicators to gates of cross-layer receptivity. Empirically, using panel data on 10,681 patents, 948 inventors and 5631 inventor-year observations drawn from the new energy field (2004–2018), it validates both the inverted U-shaped main effect and the two steepening, peak-shifting moderations with a fixed-effects negative binomial model and three robustness specifications, thereby moving multilayer network research from structural description toward mechanism-level identification. Practically, it provides network-design guidance for keeping the composite innovation network inside the optimal cross-layer integration zone, with concrete implications for managers of cross-border R&D platforms of the type discussed at the outset of this introduction.

2. Current Status of the Study

Understanding how Technology Innovation Composite Networks influence breakthrough innovation requires revisiting three strands of the literature: the role of network embeddedness in innovation, the treatment of breakthrough innovation itself, and the rise of multilayer network analysis.

Research on network embeddedness has long established that inventors' innovative behavior is shaped not only by individual technical capability but also by the structural position they occupy within knowledge and collaboration networks. Schilling [12] and Powell et al. [13] argue that innovation is strongly conditioned by network structure, and Granovetter's [14] "weak ties" argument suggests that loose, cross-boundary connections

supply inventors with novel information and technological opportunity. More recent work has operationalized these ideas with large-scale patent data and demonstrated that Technology Innovation Composite Networks can capture complex interaction patterns more precisely by integrating the knowledge layer with the collaboration layer. However, this integration has mostly remained conceptual: empirical models continue to operate layer-by-layer, and the mechanism by which information crosses from one layer to the other is rarely specified. Very recent multilayer-network studies reinforce this agenda: two-tier analyses of breakthrough innovation report inverted U-shaped network effects and show that firm digitalization reshapes them [15], and knowledge-recombination studies find the same nonlinearity in collaboration scale [16]. In parallel, a fast-growing literature documents how digitalized R&D and AI tools are beginning to restructure innovation networks and the recombination process itself [17,18], which makes the cross-layer transfer mechanism studied here increasingly timely.

Research on breakthrough innovation is similarly well developed but similarly constrained. Fleming [19] and Henderson and Clark [20] show that breakthrough innovations involve major reconfigurations of existing technologies and exert far-reaching effects on firm strategy and organization, rather than being confined to a single technological dimension. Dahlin and Behrens [21] further emphasize that breakthrough outcomes typically arise from multilevel interactions of knowledge and collaboration, in which information flow and resource integration jointly matter. These studies foreshadow the importance of cross-layer dynamics but stop short of modeling them: information flow is treated as an enabling condition rather than as a structurally coupled process that can itself be facilitated or obstructed.

The emergence of Multilayer Networks Theory [8] finally provides a formal apparatus for such coupling [19,22]. The theory treats innovation as driven by the interaction between distinct network layers, and it allows structural features—such as structural holes, degree centrality and closeness centrality—to be analyzed both within a layer and across layers. This is the theoretical lineage the present paper builds on.

Against this backdrop, three research gaps can be identified, and each maps directly onto an element of the cross-layer integration mechanism proposed here. First, most existing studies remain anchored in single-layer analysis. While Ahuja [22] verifies the importance of structural holes in a single-layer collaboration network, how structural holes in the knowledge layer translate into breakthrough outcomes through inventors' activities in the collaboration layer has not been adequately investigated. Second, the influence mechanism has often been treated as approximately linear. Fleming's [19] work reveals the relationship between search and innovation but does not fully articulate the nonlinear dynamics that a finite cognitive processing capacity would impose on cross-layer integration. Third, the moderating role of collaboration-layer centrality in multilayer settings is underexplored. Extant studies of degree centrality and closeness centrality are mostly single-layer [23,24]; they do not examine how these centralities govern the collaboration layer's capacity to receive heterogeneous information across layers, and therefore do not explain why excessively high centrality might systematically compress the inverted U-shaped main effect, pushing its peak toward lower structural-hole levels.

To confront these three gaps jointly, this paper theorizes and tests a cross-layer mechanism in which structural holes in the knowledge layer supply heterogeneous information; that information must cross into the collaboration layer to become actionable, and the effective cross-layer transfer is itself bounded—upward by the inventor's cognitive processing threshold and outward by the collaboration layer's own homogenization and localization. The remainder of the paper develops this logic in formal hypotheses and brings it to patent data from the new energy field.

3. Research Models and Hypotheses

Before developing the hypotheses, the core concepts used throughout the paper are clarified. A Technology Innovation Composite Network (TICN) is a dual-layer system that couples a knowledge network and a collaboration network: the knowledge layer links technology categories (knowledge elements) that co-occur on patents, and the collaboration layer links inventors who co-invent. Cross-layer integration is the process by which heterogeneous information identified in the knowledge layer is carried by inventors into the collaboration layer and recombined there into new technical solutions. Breakthrough innovation is innovation that reshapes a technological trajectory, operationalized here as patents in the top 5% of positive citations within each five-year window. A knowledge-layer structural hole is a gap between otherwise disconnected knowledge communities that an inventor's portfolio spans, giving access to non-redundant, heterogeneous information; following Burt, it is measured by the inverse of network constraint. Collaboration-layer centralities describe an inventor's position in the collaboration network: degree centrality is the number of direct co-inventor ties, and closeness centrality is the inverse of the average network distance to all other inventors. The optimal cross-layer integration zone is the region—defined jointly by a moderate level of structural holes and a collaboration layer that is neither over-centralized nor over-localized—within which cross-layer integration is most productive and breakthrough innovation is maximized.

Building on multilayer network theory, this paper treats the Technology Innovation Composite Network as a dual-layer system in which the knowledge network and the collaboration network are independent but coupled. The model examines how knowledge-layer structural holes drive breakthrough innovation through cross-layer integration, and how collaboration-layer centralities condition that cross-layer pathway. The overall model is shown in Figure 1.

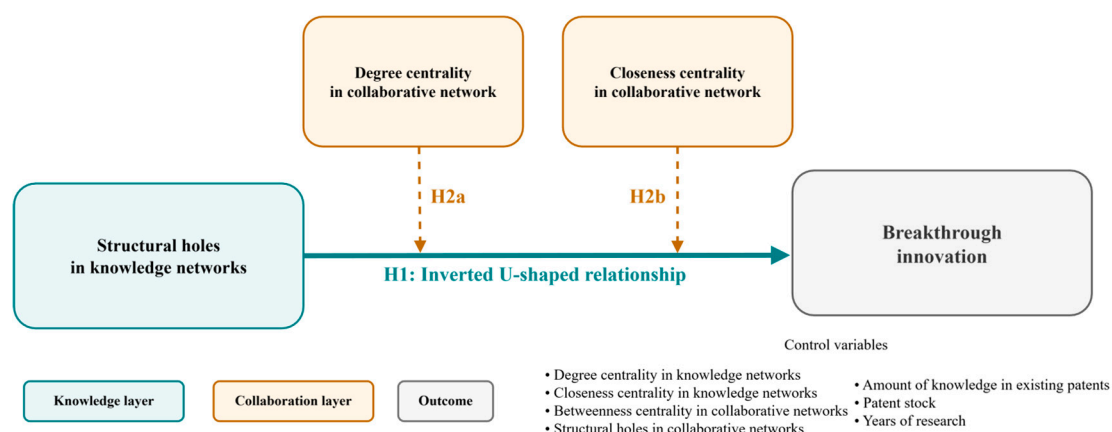


Figure 1. Diagram of the study model. Note: The solid arrow denotes the inverted U-shaped main effect from knowledge-network structural holes to breakthrough innovation; the dashed arrows denote the moderating effects of collaboration-network degree centrality and closeness centrality.

3.1. Building Knowledge Networks and Collaborative Networks

This paper uses patent data to construct the two layers. Patent data are chosen for two reasons. First, they are a core data source in innovation network research and are widely used to analyze the diffusion and impact of technological innovation [12]; they supply rich and objective information on inventors, filing time and international patent classification (IPC) [25]. Second, patent data permit controlling for time-varying variables and mitigate covariance between current and lagged quantities in ways that survey data

cannot [26], which is essential for analyzing the dynamics of a dual-layer network as it evolves.

Figure 2 illustrates the collaboration layer schematically: each node is an inventor, node size represents patent output, and edge width indicates the intensity of collaboration. Because the full empirical network contains hundreds of inventors and is too dense to render legibly, the figure highlights only representative communities and the hub inventors that bridge them; the full network exhibits pronounced local clustering together with a small-world property, so that cooperation is typically concentrated within limited communities, while a few hub inventors bridge otherwise disconnected groups. The knowledge layer is shown in the same schematic form (Figure 3), nodes represent patent outputs in specific technology categories and edge width indicates the frequency of combinations between those categories. Together, these two layers form the composite network analyzed in the remainder of the paper.

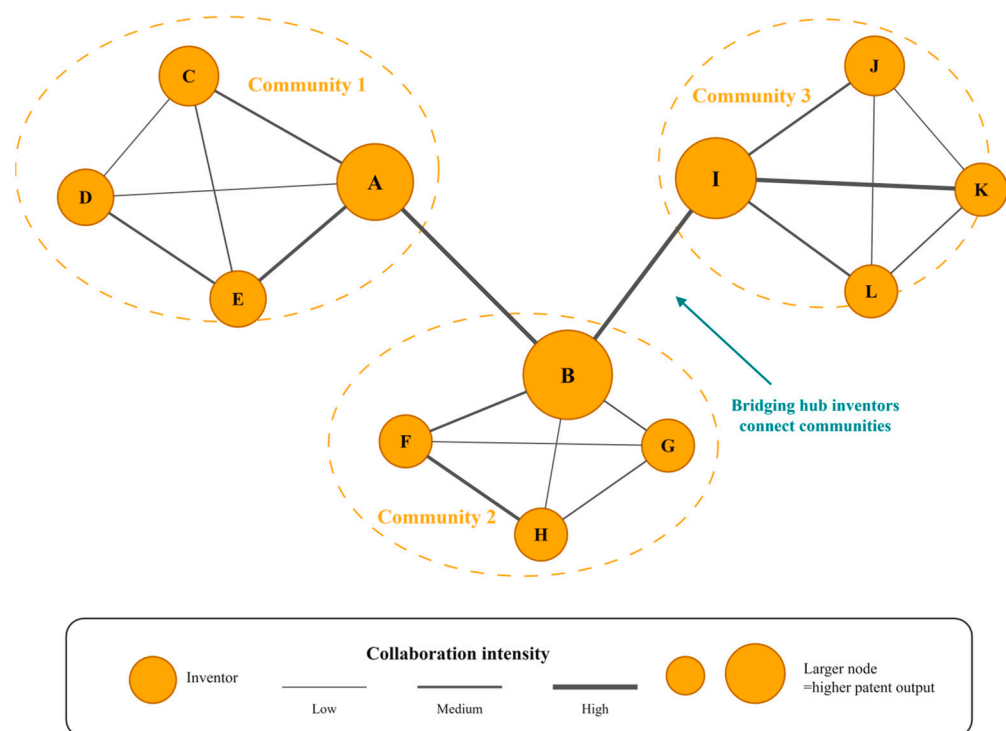


Figure 2. Schematic of the collaboration-layer structure. Note: Letters A–L denote illustrative inventors; dashed circles denote collaboration communities; solid lines denote co-invention ties; line width represents collaboration intensity; node size represents patent output; and arrows highlight hub inventors bridging otherwise disconnected communities.

Crucially, the two layers are structurally independent but functionally coupled. Figure 4 presents the two-layer network in which inventors (circles) and technology categories (squares) are displayed jointly. The figure shows that an inventor’s central position in the collaboration layer does not guarantee direct association with central technology categories in the knowledge layer; position in one layer does not automatically translate into position in the other. This structural independence is exactly why a cross-layer perspective is analytically necessary: any heterogeneous information identified in the knowledge layer must be actively carried into the collaboration layer by inventors, through their dual embedding, before it can catalyze breakthrough innovation. A single-layer analysis would collapse this transfer into an assumption and would therefore miss the mechanism of interest.

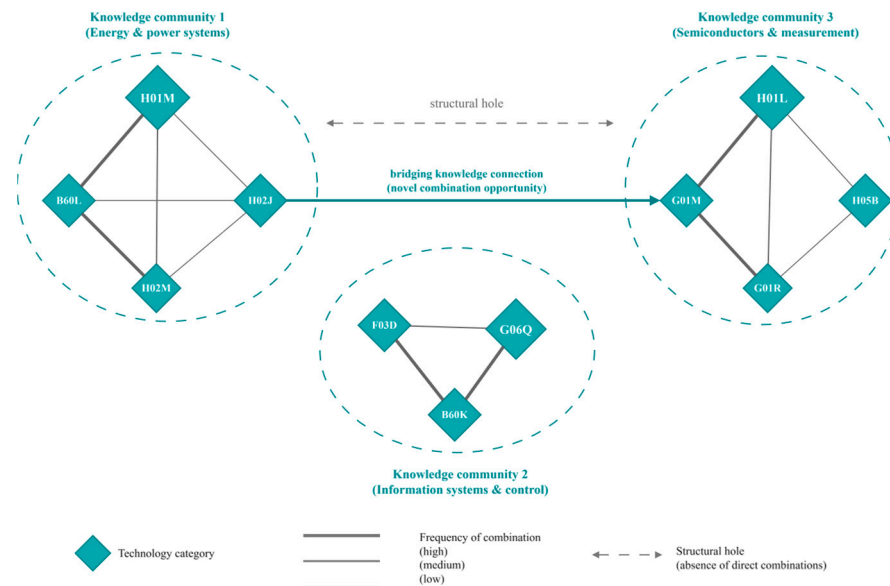


Figure 3. Schematic of the knowledge-layer structure. Note: Square patterns denote technology categories; dashed circles denote knowledge communities; solid lines denote knowledge co-occurrence ties; line width represents co-occurrence intensity; and dashed arrows indicate structural holes between otherwise disconnected knowledge communities.

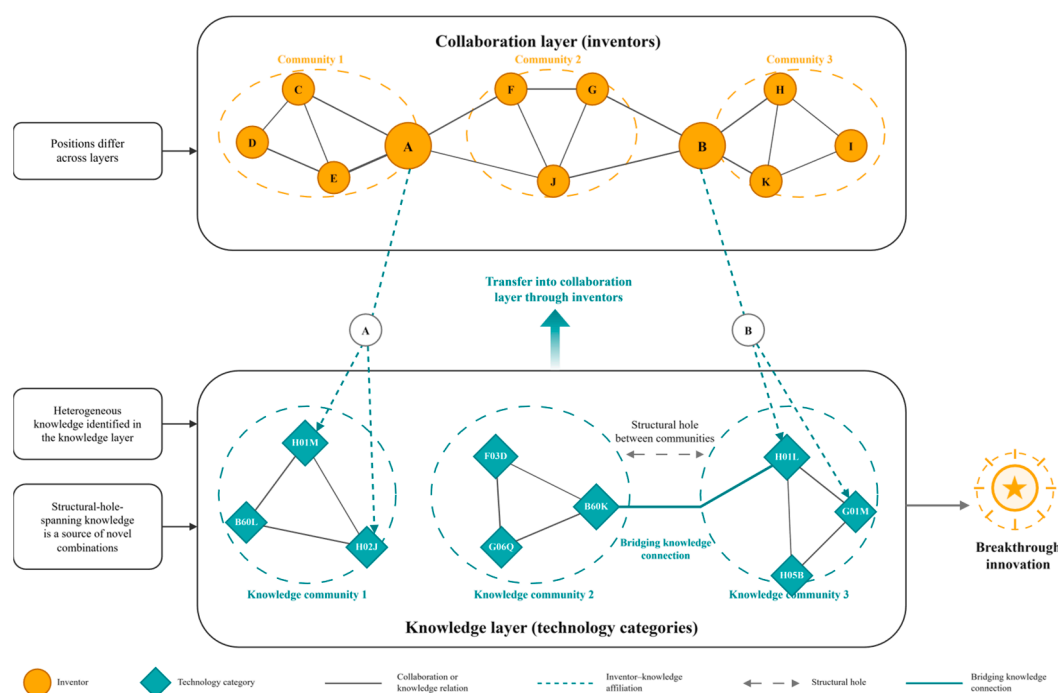


Figure 4. Two-layer network. Note: Circles denote inventors in the collaboration layer and diamonds denote technology categories in the knowledge layer. Solid lines denote within-layer ties, dashed vertical lines denote inventor–knowledge affiliations, and arrows indicate the cross-layer transfer of heterogeneous knowledge into collaborative recombination.

To operationalize this cross-layer logic, the paper adopts a lag structure in hypothesis testing: network structural features are computed over a 5-year window (year $t - 4$ to t), and breakthrough innovation outcomes are evaluated in the adjacent 5-year window (year $t + 1$ to $t + 5$). This design avoids co-occurrence, respects the time-dynamic nature of innovation, and lets the data speak to the temporal ordering implied by the cross-layer mechanism. The dataset covers 10,681 patents and 948 inventors. The number of patents and

inventors both grow exponentially between 2004 and 2018 ($R^2 = 0.968$ and $R^2 = 0.936$), while the growth of knowledge elements is relatively slow (Figure 5)—a stylized fact that is consistent with the mechanism proposed here: as inventors and their collaborations expand faster than the pool of underlying knowledge elements, recombining heterogeneous knowledge across layers becomes an increasingly binding driver of breakthrough innovation.

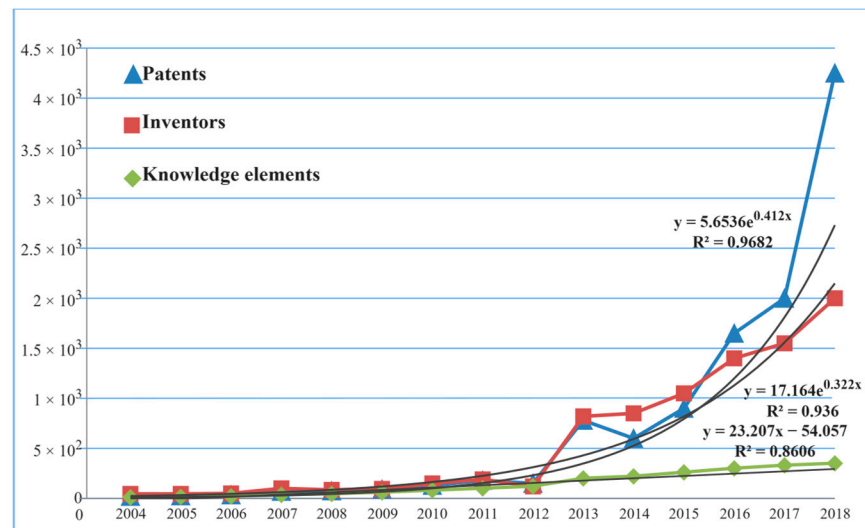


Figure 5. New energy patents at different stages. Note: The blue, red and green lines represent the yearly numbers of patents, inventors and knowledge elements, respectively; markers show observed yearly counts, and fitted curves summarize their growth trends.

3.2. Structural Holes in Knowledge Networks and Breakthrough Innovation

Within the cross-layer integration mechanism, structural holes in the knowledge network play the role of entry channels for heterogeneous information. Drawing on Burt's structural hole theory, nodes that span structural holes access novel, non-redundant information by connecting otherwise disjoint knowledge communities [7,27]. Translated into the TICN context, inventors whose knowledge portfolios span more structural holes are better positioned to detect combinatorial opportunities that cut across technological domains—opportunities that, by construction, cannot be found within a single domain [28].

The existing literature has mostly documented the positive effect of structural holes on innovation, yet has been less attentive to the bounded nature of this effect [29]. Two mechanisms bound it. The first is specific to cross-layer integration. Heterogeneous information identified in the knowledge layer does not automatically convert into breakthrough outcomes; it must be carried by inventors into the collaboration layer, where it is recombined and negotiated among collaborators [30]. The more structural holes an inventor spans, the more heterogeneous and less redundant the information that needs to be transferred, and the more demanding the coordination and integration task becomes [31]. The second mechanism is cognitive. Inventors have a finite capacity to attend to, evaluate and combine distinct streams of information. When the number of structural holes exceeds this cognitive processing threshold, additional heterogeneity is no longer absorbed into useful recombination but manifests as information overload and redundancy [32,33]. As a result, the marginal contribution of structural holes first rises and then declines, and the integration returns decay along an inverted U-shaped curve [13,34,35].

Both mechanisms operate within the multilayer network architecture. Multilayer Networks Theory emphasizes that innovation depends not only on information stocks in

any single layer but also on how information is transferred and integrated across layers [8]. Structural holes thus affect breakthrough innovation precisely through cross-layer integration: they widen the set of heterogeneous inputs inventors can draw upon, but only up to the point at which the inventor's own processing capacity ceases to scale with that input. Beyond this point, more structural holes generate more coordination burden than combinational gain [13,30].

Based on this logic, the following hypothesis is proposed:

H1. *There is an inverted U-shaped relationship between structural holes in the knowledge network and an inventor's breakthrough innovation.*

The proposed inverted U-shape is not simply a refinement of structural hole theory; it is also a direct consequence of embedding structural holes within a cross-layer integration mechanism [7,8,33,34]. Understanding where the curve peaks is substantively meaningful, because it identifies the intensity of heterogeneous-information access at which cross-layer integration is most productive—the upper boundary of the optimal integration zone from the knowledge-layer side [8,13,30].

3.3. The Moderating Role of Degree Centrality and Closeness Centrality in Collaborative Networks

Whether the heterogeneous information supplied by structural holes can actually cross layers depends on the state of the collaboration network itself. In the TICN, inventors are simultaneously embedded in the knowledge layer and the collaboration layer (Figure 6), and the collaboration layer is the space in which cross-domain information is translated into joint action, technical solutions and ultimately breakthrough outputs [11,36,37].

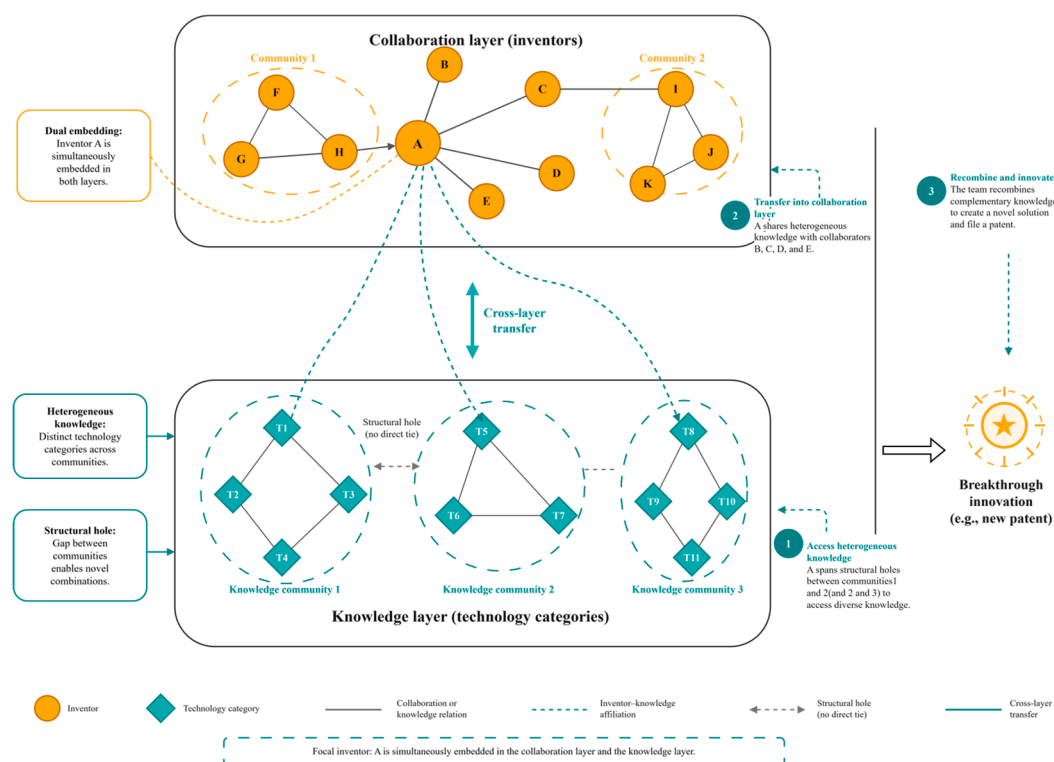


Figure 6. Inventor's simultaneously embedded two-layer network. Note: The upper panel represents the collaboration layer and the lower panel represents the knowledge layer. Dashed vertical links indicate the inventor's simultaneous embedding across layers, while arrows denote the cross-layer transfer and recombination pathway leading to breakthrough innovation.

Collaboration-network position, however, is a double-edged construct. It can accelerate information flow and resource access, but when a single node becomes too central—either in the sense of directly connecting to many others (degree centrality) or in the sense of being uniformly close to all others (closeness centrality)—the collaboration layer loses its structural diversity [7]. Two opposing scenarios make this concrete. Inventors with very high degree centrality obtain resources quickly, but information becomes overly concentrated, leading to duplication, redundancy and a narrowing of innovation paths [38]. Inventors with very high closeness centrality transfer information rapidly, but mostly among neighboring nodes, so that the network exchange becomes localized and loses heterogeneity [39,40].

From the standpoint of the cross-layer integration mechanism, both patterns have the same consequence: they degrade the collaboration layer's capacity to receive heterogeneous information from the knowledge layer. When the collaboration layer is already homogenized (high degree centrality) or localized (high closeness centrality), heterogeneous inputs brought in by structural-hole-spanning inventors are absorbed into an environment that resembles what those inputs were supposed to disrupt. The cross-layer transfer, in other words, still physically takes place, but it no longer triggers the recombination that breakthrough innovation requires. The resulting moderating effect is therefore not merely “weakening” in a vague sense; it specifically reshapes the inverted U-shaped curve: the collaboration layer reaches its breakthrough peak at a lower intensity of structural holes and the returns decline more steeply beyond that peak, so the range of structural holes the layer can productively absorb narrows [40,41].

Degree centrality captures the direct connectivity of an inventor in the collaboration network and is a standard indicator of access to information and resources [23]. Inventors at the center of the information flow can assemble resources quickly and thus, in moderate amounts, complement the cross-layer integration mechanism. Beyond a moderate level, however, excessive degree centrality concentrates information, exposes inventors to overload, and homogenizes the local collaboration environment; the heterogeneous information flowing in from the knowledge layer can no longer find the distinct collaborators needed to recombine it. Consequently, the collaboration layer loses part of its across-layer receptive capacity, and the inverted U-shaped main effect peaks at a lower structural-hole level and declines more steeply beyond it. Hence:

H2a. *Degree centrality in the collaborative network negatively moderates the inverted U-shaped relationship between structural holes in the knowledge network and breakthrough innovation.*

Closeness centrality captures the ease with which an inventor reaches other nodes through short network paths [14]. High closeness centrality accelerates the diffusion of ideas in the inventor's immediate surroundings, but when closeness centrality becomes excessive, collaboration crystallizes around neighboring nodes and the inventor's effective reach contracts into a local clique. The collaboration layer then becomes localized: cross-domain heterogeneous information arriving from the knowledge layer can no longer diffuse across distant parts of the collaboration network, and the full advantage of the structural-hole-opened channel is lost. As a result, the inverted U-shaped curve steepens and its peak shifts toward lower structural-hole levels, implying the same narrowing of the productive window as in H2a. Hence:

H2b. *Closeness centrality in the collaborative network negatively moderates the inverted U-shaped relationship between structural holes in the knowledge network and breakthrough innovation.*

Taken together, H2a and H2b articulate the second boundary condition of the cross-layer integration mechanism: the collaboration layer must itself remain heterogeneous

and globally connected enough to receive, distribute and recombine the heterogeneous information produced by structural holes. Together with H1, they identify an optimal cross-layer integration zone—a region defined by moderate structural holes in the knowledge layer and moderate (not excessive) centrality in the collaboration layer—within which breakthrough innovation reaches its maximum.

4. Study Design

4.1. Data Collection

Hypothesis testing is conducted using panel data on patents published in the new energy field between 2004 and 2018. The panel comprises 5631 inventor-year observations involving 948 inventors (57 observations carrying logically impossible negative research-tenure values were removed during data cleaning) [42]. To avoid statistical omissions caused by the time lag between patent application and publication, 2018 is chosen as the end year of data collection [43].

Data-processing workflow. Because the validity of the findings depends on how the two layers are built, the full data-processing workflow is reported here so that it can be reproduced from an equivalent patent extract. (i) Patent retrieval and inclusion. New-energy patents granted between 2004 and 2018 were retrieved from the Derwent Patent Database using the database's new-energy technology classification together with the corresponding International Patent Classification (IPC) codes; records missing inventor, application-year or IPC information were discarded, and only granted patents were retained so that positive-citation counts are well defined. (ii) Inventor disambiguation. Inventor identities were resolved by combining inventor name strings with assignee, co-inventor and IPC-field information, so that variant spellings of the same inventor are merged and distinct inventors who share a name are separated; this procedure yields the 948 unique inventors analyzed. (iii) Patent assignment and inventor-year observations. Each patent was attributed to all of its disambiguated inventors, and the panel was organized as inventor-year observations on rolling five-year windows, with network features measured over the five years up to and including year t and breakthrough outcomes measured over the following five years; 57 observations carrying logically impossible negative research-tenure values were removed during cleaning, leaving 5631 inventor-year observations. (iv) IPC processing and knowledge elements. Knowledge elements are defined at the IPC subclass level, and two subclasses are linked in the knowledge layer whenever they co-occur on the same patent. (v) Edge weighting and network boundary. In the collaboration layer, inventors are nodes and an edge connects two inventors who have co-invented at least one patent, weighted by the number of co-invented patents; in the knowledge layer, IPC subclasses are nodes and edges are weighted by their co-occurrence frequency. Both layers are rebuilt for each five-year window, and the network boundary for a window comprises the inventors and the IPC subclasses that appear on the patents granted within that window. (vi) Matrices. The workflow yields, for each window, an inventor-year matrix together with two weighted adjacency matrices—an inventor-by-inventor collaboration matrix and an IPC-by-IPC knowledge matrix—from which all structural-hole and centrality measures were computed in UCINET 6.0. Because the underlying records are licensed from the Derwent database and cannot be redistributed, the raw patent data cannot be shared; the search strategy, inclusion criteria, disambiguation rules, IPC-coding scheme and matrix-construction steps set out above are nevertheless reported in enough detail to reconstruct the inventor-year and network matrices from an equivalent extract.

The new energy field is a well-suited empirical setting for the cross-layer integration mechanism for three reasons. First, new energy technologies are central to addressing global energy shortages and environmental pollution caused by traditional energy use [44], so breakthrough innovation in this field carries intrinsic substantive importance.

Second, the rapid development of new energy technologies over the past two decades has produced a large volume of patent records in the Derwent Patent Database, supplying the sample density required for a dual-layer analysis [45,46]. Third—and most directly relevant to the mechanism studied here—more than 64.80% of new energy patents integrate multiple knowledge elements and exhibit rich collaboration relationships [47]. This means that the sample itself is dominated by patents whose creation requires combining heterogeneous knowledge across technological domains through collaboration, which is precisely the cross-layer integration process the theoretical model seeks to capture.

4.2. Measurement Methods

Breakthrough Innovation. Breakthrough patents are defined as those ranking in the top 5% of positive citation counts within each 5-year time window [48]. The number of breakthrough patents is used as the measure of breakthrough innovation [48,49]. In operational terms, the cumulative positive-citation count accrued by each patent between its grant date and the download date is retrieved from the Derwent patent database [50]. The list of breakthrough patents M for a given observation window is then constructed from the top 5% of patents identified by positive citations from year $t + 1$ to year $t + 5$, and the number of patents under each inventor's name that appear on this list is counted. Equation (1) formalizes this lag structure:

$$\text{Breakthrough Innovation}_{k,t-4 \sim t} = \text{Number of Patents in } M_{k,t+1 \sim t+5} \quad (1)$$

where $\text{Breakthrough Innovation}_{k,t-4 \sim t}$ denotes the breakthrough innovation of inventor k during the five-year window from year $t - 4$ to year t , and $\text{Number of Patents in } M_{k,t+1 \sim t+5}$ is the count of inventor k 's patents on the breakthrough list for the subsequent window. This design ensures that the network features used as predictors are measured before the breakthrough outcomes they are hypothesized to drive.

The top-5% threshold is chosen to balance selectivity against sample density in the new-energy setting. New-energy patenting is voluminous and its positive-citation distribution is highly right-skewed, so a top-5% cut within each window is restrictive enough to isolate the small set of patents that genuinely reshape a technological trajectory, yet still leaves enough breakthrough events per window to estimate the inverted U-shape and its moderators with reasonable precision. A much stricter cut (for example, the top 1%) yields very few positive cases in the smaller windows and inflates the variance of the curvature estimate, whereas a looser cut (for example, the top 10%) dilutes the breakthrough category with incremental patents and weakens the contrast the construct is meant to capture. To show that the conclusions do not hinge on this choice, the operationalization is varied in two directions in Section 5.3: a stricter top-1% threshold (Sensitivity A) and a field-and-year-standardized citation measure that removes the cross-field and over-time differences in citation intensity characteristic of fast-moving new-energy subfields (Sensitivity B). Both reproduce the inverted U-shaped main effect and the two moderations, indicating that the dependent variable is not an artifact of the 5% cut-off.

Structural holes in knowledge networks. Burt's constraint index (C_{ij}) is used to measure the extent to which knowledge elements in an inventor's portfolio occupy structural holes in the knowledge network [51]. A knowledge network is constructed at each analysis stage, and UCINET 6.0 is used to compute the constraint index for each knowledge element [52]. The constraint index is:

$$C_{ij} = (p_{ij} + \sum_q p_{iq}p_{qj})^2, i \neq j \neq q \quad (2)$$

and the structural hole metric is:

$$S_i = 2 - \sum_j \left(p_{ij} + \sum_{q, q \neq i, q \neq j} p_{iq} p_{qj} \right)^2 \quad (3)$$

where p_{ij} , p_{iq} , and p_{qj} denote the ratios of direct connections between knowledge elements i and j , i and q , and q and j , respectively, to the total connections ($i \neq j \neq q$). As Burt shows, a high network constraint implies fewer structural holes, so network constraint is negatively correlated with spanning structural holes [23,51]. The structural holes of each inventor's knowledge network (denoted KSH) are computed as the average level of structural holes of the knowledge elements in that inventor's portfolio [23]. Averaging across an inventor's knowledge-element set yields a portfolio-level measure of the heterogeneous-information access channels opened to that inventor.

Degree centrality in the collaborative network. Degree centrality captures the number of inventors a focal inventor is directly connected to in the collaboration network [53,54]. A collaboration network is constructed at each analysis stage, with inventors as nodes and co-invention ties as edges. UCINET 6.0 is then used to compute degree centrality according to:

$$C_i = \frac{c_i}{N-1} = \frac{\sum_{j \in G} a_{ij}}{N-1} \quad (4)$$

where c_i is the degree of inventor i , N is the total number of inventors in collaboration network G , and the adjacency element $a_{ij} = 1$ if inventors i and j have collaborated and 0 otherwise (with a_{ii} set to 1 for convenience) [53,54]. The resulting measure is denoted CDC.

Closeness centrality in the collaborative network. Closeness centrality is measured through the sum of shortest-path lengths between a node and all other nodes in the network, and is normalized as in Equation (5) [55]:

$$C_c^c(n_i) = \frac{(g-1)}{\sum_{j=1}^g d(n_i, n_j)} \quad (5)$$

where $C_c(n_i)$ is the closeness centrality of node n_i in a network of g nodes, and $d(n_i, n_j)$ is the shortest-path length between n_i and n_j . The resulting measure is denoted CCC.

Control variables. Six controls are included to absorb alternative explanations for breakthrough innovation.

Patent stock controls for inventors' past technological creativity and absorptive capacity, measured by the number of patents an inventor received in the five years prior to entering the empirical sample [56,57].

Betweenness centrality in the collaborative network (CBC) controls for an inventor's bridging role in the collaboration network, normalized as in Equation (6) [58]:

$$BC(v)_{\text{norm}} = \frac{2}{(N-1)(N-2)} \sum_{v \neq i \neq j} \frac{g_{ij}(v)}{g_{ij}} \quad (6)$$

where g_{ij} and $g_{ij}(v)$ are the numbers of shortest paths between i and j that do not pass through v and that do pass through v , respectively. Controlling for betweenness isolates the effect of degree and closeness centrality from the bridging function of the inventor.

Years of research is defined by the interval between the current year and the year of the inventor's first patent application, capturing the accumulated experience that influences the ability to produce breakthrough innovation [59,60].

Closeness centrality in knowledge networks (KCC) is computed as in Equation (7) to control for the within-knowledge-layer efficiency of information diffusion [55]:

$$C_c^K(n_i) = \frac{(g-1)}{\sum_{j=1}^g d(n_i, n_j)} \quad (7)$$

Amount of knowledge in existing patents is the sum of knowledge elements involved in an inventor's patents during the observation window, controlling for the breadth and heterogeneity of the inventor's existing knowledge combinations [47,61].

Degree centrality in knowledge networks (KDC) is computed as in Equation (8), capturing the within-knowledge-layer influence of knowledge elements in the inventor's portfolio [53,54]:

$$C_{D(n_i)} = \frac{d(n_i)}{g-1} = \frac{\sum_{j=1}^g z_{ij}}{g-1} = \frac{\sum_{j=1}^g z_{ji}}{g-1} \quad (8)$$

where $z_{ij} = 1$ indicates an association between knowledge elements i and j (z_{ii} set to 1 for simplicity). Jointly, the three knowledge-layer controls (KDC, KCC, amount of knowledge in existing patents) and the collaboration-layer control (CBC) ensure that the estimated effects of KSH, CDC and CCC reflect the cross-layer integration mechanism rather than generic within-layer position effects.

For compactness, the following symbols are used in Equation (17), in all tables, and in the remainder of the paper: KSH denotes knowledge-network structural holes (Equations (2) and (3)); CDC, CCC and CBC denote degree, closeness and betweenness centrality in the collaboration network (Equations (4)–(6)); and KDC and KCC denote degree and closeness centrality in the knowledge network (Equations (7) and (8)). A squared term is written, for example, KSH^2 , and a moderation as $KSH^2 \times CDC$.

Model Selection. The empirical model evaluates how multilevel network features during year $t - 4$ to year t affect breakthrough innovation during year $t + 1$ to year $t + 5$. Because breakthrough innovation is a count variable, Poisson regression with maximum likelihood estimation is a natural starting point [62]; its count-data density is given by

$$p(Y_i = y_i) = p(y_i) = \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!} \quad i = 1, 2, 3, \dots, n \quad (9)$$

with

$$\mu_i = E(Y_i) = v_i \left[e^{x_i^j \beta} \right] = v_i \left[e^{\sum_{j=1}^k x_{ij} \beta_j} \right] \quad i = 1, 2, 3, \dots, n \quad (10)$$

However, the Poisson specification requires equidispersion. Breakthrough innovation across inventors is typically highly skewed, so the conditional variance is expected to exceed the conditional mean—i.e., over-dispersion is anticipated—making a negative binomial specification preferable to Poisson [63,64] (The empirical extent of over-dispersion in this sample is reported in Section 5.1). In addition, two distinct data-generating regimes are expected to produce excess zeros: inventors who filed no patents during the observation window, and inventors who filed patents but none entered the top-5% breakthrough list. A zero-inflated negative binomial (ZINB) model [65,66] is therefore used to frame the modeling strategy: the main specification is a fixed-effects negative binomial model for the over-dispersed count process, while a zero-inflated specification is used in the robustness checks to address the excess-zero mechanism (The actual share of zeros in the sample is reported in Section 5.1). The ZINB density is:

$$P_r(Y = y) = \begin{cases} p + (1-p) \left(1 + \frac{\mu}{\tau}\right)^{-\tau} & y = 0 \\ (1-p) \frac{\Gamma(y+\tau)}{y! \Gamma(\tau)} \left(1 + \frac{\mu}{\tau}\right)^{-\tau} \left(1 + \frac{\tau}{\mu}\right)^{-y} & y = 1, 2, \dots \end{cases} \quad (11)$$

where Γ is the gamma function, τ is the over-dispersion parameter, and μ is the expected mean. The conditional mean and variance are:

$$E(Y) = (1-p)\mu \quad (12)$$

$$\text{var}(Y) = (1-p)\mu \left(1 + p\mu + \frac{\mu}{\tau}\right) \quad (13)$$

The ZINB nests a zero-inflated Poisson as $\tau \rightarrow \infty$ and a negative binomial as $p \rightarrow 0$. The covariate structure is specified as:

$$\ln(\mu_i) = x_i' \beta \quad (14)$$

$$\log \text{it}(p_i) = \ln\left(\frac{p_i}{1-p_i}\right) z_i' \gamma, (i = 1, 2, \dots, n) \quad (15)$$

and the log-likelihood is:

$$\begin{aligned} L_z(\beta, \gamma, \tau, y, X, Z) = & \sum_{i=1} \left(1 + e^{z_i' \gamma}\right) + \sum_{i: y_i=0} \log \left(e^{z_i' \gamma} + \left(\frac{e^{x_i' \beta} + \tau}{\tau} \right)^{-\tau} \right) \\ & + \sum_{i: y_i>0} \left(\tau \log \left(\frac{e^{x_i' \beta} + \tau}{\tau} \right) + y_i \log \left(1 + e^{-x_i' \beta \tau} \right) \right) \\ & + \sum_{i: y_i>0} (\log \Gamma(\tau) + \log \Gamma(1 + y_i) - \log \Gamma(\tau + y_i)) \end{aligned} \quad (16)$$

with $X = (x_1, \dots, x_n)$, $Z = (z_1, \dots, z_n)$. Because the panel involves repeated observations on the same inventors, time-invariant inventor heterogeneity—innate creativity, intrinsic preference for breakthrough versus incremental work, organizational affiliation effects—is likely correlated with network position and would bias a pooled or random-effects estimator. A fixed-effects negative binomial specification is adopted as the main model to absorb this heterogeneity while accommodating over-dispersion, and the zero-inflated variant is retained for the robustness checks in Section 5.3. In implementation, the main model is the Hausman–Hall–Griliches conditional fixed-effects negative binomial estimator, with the inventor as the panel unit and time-window indicators included to absorb period-specific shocks; the Poisson and zero-inflated Poisson specifications in Section 5.3 corroborate the estimates and guard against the known limitations of the conditional negative-binomial estimator. In the zero-inflated specifications, the inflation (zero-generating) equation uses the same covariates as the count equation, so that both the structural-zero process and the breakthrough-count process are modeled with the full set of network and control variables (The Hausman test comparing fixed and random effects is reported in Section 5.2). The general form of the estimated equation is:

$$\begin{aligned} & BL_{(i,t+1 \sim t+5)} \\ & = f(KSH, KSH^2, KSH \times CDC, KSH^2 \times CDC, KSH \times CCC, KSH^2 \\ & \times CCC, KDC, KCC, CBC, \text{amount of knowledge in existing patents, years of research, patent stock}) \end{aligned} \quad (17)$$

The squared term of KSH tests the curvature predicted by H1, while the interaction terms between KSH² and the two collaboration-layer centralities test the moderations predicted by H2a and H2b—a steeper, earlier-peaking inverted U as centrality rises—directly operationalizing the cross-layer integration mechanism. To verify robustness, three alternative specifications are also estimated: negative binomial regression with robust standard errors [67,68], Poisson regression [62] and zero-inflated Poisson regression. Robustness results are reported in Section 5.3.

5. Data Analysis and Results

Data processing, statistical analysis and visualization are conducted with Stata 17.0, Python 3.9.13 and MATLAB R2022a.

5.1. Descriptive Statistics

Table 1 reports descriptive statistics and bivariate correlations for all variables. The sample covers 10,681 patents produced by 948 inventors between 2004 and 2018, yielding 5631 inventor-year observations after data cleaning. The bivariate correlation between structural holes (KSH) and breakthrough innovation is negative and highly significant ($r = -0.276$, $p < 0.001$). This negative sign does not contradict the inverted-U hypothesis: because the sample mean of structural holes (1.803) already lies beyond the inverted-U turning point identified in Section 5.2 (1.438), most observations fall on the declining segment

of the curve, so the simple bivariate correlation is negative even though the underlying relationship is curvilinear. All non-DV variance inflation factors, computed in a separate diagnostic regression, are below 10 (maximum 5.28), so multicollinearity is not a concern. In addition, the dependent variable is over-dispersed (variance/mean = 2.58) and 50.72% of inventor-year observations record zero breakthrough patents, justifying the negative-binomial count specification and the zero-inflated robustness checks adopted in Sections 4.2 and 5.3.

Table 1. Descriptive statistics and variable correlations.

Variable	Mean	SD	Min	Max	VIF	1	2	3	4	5	6	7	8	9	10
Breakthrough innovation	0.841	1.474	0	23	—										
Structural holes	1.803	0.346	0.789	2.000	4.60	−0.276 ***									
Degree centrality	0.085	0.299	0	4.661	1.47	0.173 ***	−0.509 ***								
Closeness centrality	0.095	0.183	0	0.549	4.36	0.333 ***	−0.846 ***	0.527 ***							
Amount of knowledge in existing patents	1.328	4.756	0	186	5.28	0.464 ***	−0.441 ***	0.248 ***	0.450 ***						
Patent stock	0.386	2.185	0	120	3.91	0.348 ***	−0.291 ***	0.220 ***	0.297 ***	0.852 ***					
Years of research	23.561	5.730	0	41	1.09	0.265 ***	−0.262 ***	0.137 ***	0.268 ***	0.184 ***	0.143 ***				
Knowledge degree centrality	0.111	0.291	0	2.613	2.08	0.168 ***	−0.500 ***	0.344 ***	0.577 ***	0.311 ***	0.192 ***	0.152 ***			
Knowledge closeness centrality	0.514	1.022	0	3.702	3.75	0.225 ***	−0.774 ***	0.495 ***	0.746 ***	0.363 ***	0.247 ***	0.222 ***	0.693 ***		
Betweenness centrality	157.10	1082.845	0	22,360	1.59	0.572 ***	−0.223 ***	0.159 ***	0.298 ***	0.582 ***	0.461 ***	0.143 ***	0.139 ***	0.160 ***	
Collaboration structural holes	1.835	0.316	0.875	2.000	—	−0.120 ***	0.849 ***	−0.263 ***	−0.623 ***	−0.305 ***	−0.188 ***	−0.198 ***	−0.566 ***	−0.778 ***	−0.021

Note: N = 5631. Bivariate Pearson correlations are reported; VIF values are from a separate diagnostic regression. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (2-tailed).

5.2. Fixed-Effects Negative Binomial Regression

A Hausman test comparing fixed-effects and random-effects panel specifications rejects the null of consistent random effects ($p < 0.001$), supporting the fixed-effects choice made in Section 4.2.

Table 2 reports the fixed-effects negative binomial (FE-NB) results, estimated with time-window fixed effects and with all continuous predictors standardized. Model 1 contains only control variables. Model 2 adds structural holes (KSH), its squared term (KSH^2), and the two moderators, degree centrality (CDC) and closeness centrality (CCC). Models 3 and 4 are full models that additionally include, respectively, the degree-centrality interactions ($KSH \times CDC$ and $KSH^2 \times CDC$) and the closeness-centrality interactions ($KSH \times CCC$ and $KSH^2 \times CCC$).

Model 1 shows that breakthrough innovation is positively associated with the amount of knowledge in existing patents ($\beta = 0.147$, $p < 0.001$), knowledge-network closeness centrality KCC ($\beta = 0.120$, $p < 0.001$), collaboration-network betweenness centrality CBC ($\beta = 0.126$, $p < 0.001$) and years of research ($\beta = 0.238$, $p < 0.001$), consistent with the baseline literature.

Model 2 tests H1. The coefficient on structural holes (KSH) is positive and significant ($\beta = 1.064$, $p < 0.01$), while the coefficient on its squared term (KSH^2) is negative and significant ($\beta = -0.128$, $p < 0.01$). The joint significance of these two coefficients confirms an inverted U-shaped relationship between knowledge-layer structural holes and breakthrough innovation [66], supporting H1. Substantively, the inverted U confirms the bounded nature of the cross-layer integration mechanism from the knowledge-layer side:

structural holes raise heterogeneous-information access and thus breakthrough potential up to a certain intensity, after which further structural holes trigger information overload and redundancy and the integration returns decay.

Models 3 and 4 test H2a and H2b, with moderator variables and their interaction terms standardized to reduce multicollinearity [67]. In Model 3, the interaction between degree centrality and the squared structural-hole term ($KSH^2 \times CDC$) is significantly negative ($\beta = -0.138$, $p < 0.05$), indicating that higher degree centrality in the collaboration layer steepens the inverted U-shaped main effect and shifts its peak toward lower structural-hole levels, narrowing the productive window. This is consistent with H2a: as the collaboration layer becomes overly centralized, it homogenizes internally and loses receptive capacity for the heterogeneous information produced by structural holes, so the inverted U curve steepens and its peak shifts to a lower structural-hole level. In Model 4, the interaction between closeness centrality and the squared structural-hole term ($KSH^2 \times CCC$) is also significantly negative ($\beta = -0.243$, $p < 0.001$), supporting H2b: as closeness centrality rises, collaboration becomes localized and again loses the capacity to channel heterogeneous information across layers, further steepening the curve and shifting its peak toward a lower structural-hole level [68,69]. An earlier conference version of this study described these centrality moderations as positive; that characterization followed the sign of the linear interaction term, whereas the negative coefficient on the squared interaction term ($KSH^2 \times CDC$ and $KSH^2 \times CCC$)—which Section 5.3 shows to be robust across estimators—indicates that higher centrality steepens the inverted-U and shifts its turning point toward lower structural-hole levels. The present study, therefore, adopts the more precise interpretation that collaboration-layer centrality compresses the structural-hole window over which cross-layer integration remains productive.

Table 2. Regression model estimation results.

Variable	Model 1	Model 2	Model 3	Model 4
Methods of analysis	Fixed-effects negative binomial regression model			
Structural holes (KSH)		1.064 ** (2.78)	1.228 ** (2.92)	1.632 ** (2.78)
Structural holes ² (KSH ²)		−0.128 ** (−2.87)	−0.154 ** (−3.15)	−0.215 ** (−3.24)
Degree centrality (CDC)		0.151 *** (9.70)	−1.260 (−1.94)	0.140 *** (9.07)
Closeness centrality (CCC)		−0.015 (−0.47)	−0.000 (−0.01)	−2.847 *** (−3.52)
KSH × CDC			0.889 * (2.35)	
KSH ² × CDC			−0.138 * (−2.49)	
KSH × CCC				1.687 *** (4.00)
KSH ² × CCC				−0.243 *** (−4.46)
Knowledge degree centrality (KDC)	−0.029 (−1.46)	−0.055 * (−2.42)	−0.041 (−1.77)	−0.007 (−0.28)
Knowledge closeness centrality (KCC)	0.120 *** (5.82)	−0.026 (−0.78)	−0.004 (−0.13)	−0.021 (−0.60)
Betweenness centrality (CBC)	0.126 *** (9.49)	0.112 *** (9.51)	0.112 *** (9.50)	0.104 *** (9.06)
Amount of knowledge in existing patents	0.147 ***	0.119 ***	0.121 ***	0.110 ***

	(5.94)	(4.87)	(4.92)	(4.56)
Patent stock	−0.062 **	−0.055 **	−0.057 **	−0.052 **
	(−2.89)	(−2.74)	(−2.84)	(−2.68)
Years of research	0.238 ***	0.224 ***	0.226 ***	0.222 ***
	(12.74)	(12.09)	(12.18)	(12.04)
Constant	−1.488 ***	−3.468 ***	−3.780 ***	−4.463 ***
	(−19.76)	(−4.62)	(−4.50)	(−3.66)
Time-window fixed effects	Yes	Yes	Yes	Yes
Observations	5631	5631	5631	5631
Log likelihood	−6090.886	−6038.597	−6034.579	−6022.629
Pseudo R ²	0.148	0.155	0.156	0.157

Note: z-statistics in parentheses; all continuous predictors are standardized; time-window fixed effects included. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

To visualize the two moderations, the inverted U-shaped curve implied by the main interaction model is evaluated at low, medium and high levels of each moderator—the 25th, 50th and 90th percentiles of that moderator’s non-zero distribution—following Aiken’s (1991) conditional-curve procedure [69]. Figure 7 plots the three curves for degree centrality. As degree centrality rises, the inverted U-shaped curve becomes more sharply concave and its turning point shifts toward lower structural-hole levels: the conditional quadratic term grows steadily more negative (from -0.119 under low to -0.565 under high degree centrality) and the breakthrough-maximizing level of structural holes falls from 1.459 to 1.188. A more degree-centralized collaboration layer therefore reaches its breakthrough peak at a lower intensity of structural holes, and the returns decline more steeply beyond that peak, so the band of structural holes the layer can productively absorb narrows—the collaboration layer saturates sooner, consistent with the reduced cross-layer receptivity predicted by H2a. Figure 8 plots the analogous curves for closeness centrality and displays the same pattern even more strongly: the conditional quadratic term steepens from -0.354 to -0.805 and the turning point moves from 1.268 to 1.229 as closeness centrality rises, so a more localized collaboration layer likewise saturates at a lower structural-hole level and declines more steeply beyond it, supporting H2b.

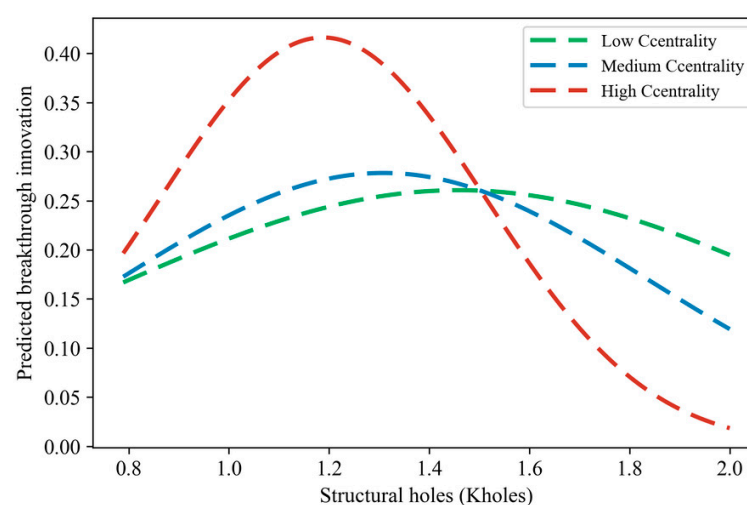


Figure 7. Moderating effects of degree centrality in collaborative networks.

Conditional inverted-U curves for degree centrality (conditional linear and quadratic coefficients, turning point, and its 95% confidence interval, on the structural-holes scale):

- High degree centrality (CDC ≈ 0.98): linear coefficient = +3.883 ($p < 0.001$); quadratic coefficient = -0.565 ($p < 0.01$); turning point = 1.188, 95% CI [1.116, 1.259].
- Medium degree centrality (CDC ≈ 0.21): linear coefficient = +1.603 ($p < 0.001$); quadratic coefficient = -0.212 ($p < 0.001$); turning point = 1.308, 95% CI [1.197, 1.420].
- Low degree centrality (CDC ≈ 0.01): linear coefficient = +1.005 ($p < 0.05$); quadratic coefficient = -0.119 ($p < 0.05$); turning point = 1.459, 95% CI [1.350, 1.567].

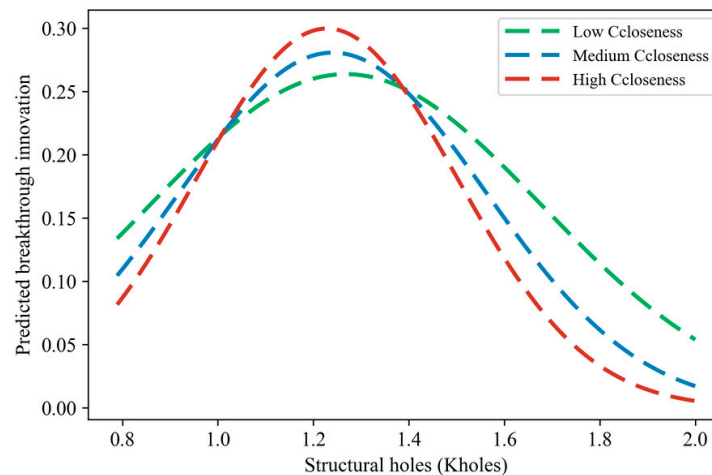


Figure 8. Moderating effects of closeness centrality in collaborative networks.

Conditional inverted-U curves for closeness centrality (conditional linear and quadratic coefficients, turning point, and its 95% confidence interval, on the structural-holes scale):

- High closeness centrality (CCC ≈ 0.54): linear coefficient = +5.722 ($p < 0.001$); quadratic coefficient = -0.805 ($p < 0.001$); turning point = 1.229, 95% CI [1.189, 1.268].
- Medium closeness centrality (CCC ≈ 0.37): linear coefficient = +4.159 ($p < 0.001$); quadratic coefficient = -0.580 ($p < 0.001$); turning point = 1.241, 95% CI [1.196, 1.286].
- Low closeness centrality (CCC ≈ 0.20): linear coefficient = +2.596 ($p < 0.001$); quadratic coefficient = -0.354 ($p < 0.001$); turning point = 1.268, 95% CI [1.180, 1.356].

Taken together, Figures 7 and 8 give a geometric image of the optimal cross-layer integration zone: breakthrough innovation is maximized when the intensity of structural holes is matched to the state of the collaboration layer—a moderate level of structural holes when the collaboration layer is diverse and globally connected, and a progressively lower level as the layer becomes more centralized or localized and its absorptive range contracts. Departures from this matched region move inventors off the peak through two distinct mechanisms. Too few or too many structural holes decay the main effect along the curve; rising degree or closeness centrality steepens the curve and shifts its peak toward lower structural-hole levels, narrowing the band of structural holes over which cross-layer integration remains productive. Both effects are theoretically implied by the cross-layer integration mechanism, and the data jointly confirm them.

Turning point and curvature diagnostics. Although Models 2–4 establish a negative quadratic term, an inverted-U interpretation is warranted only if the curve turns within the observed data and rises, and then falls across that range[70]. Table 3 reports the relevant diagnostics. The estimated turning point on the structural-holes scale is 1.438, and both its Delta-method 95% confidence interval [1.347, 1.529] and the more conservative Fieller interval [1.275, 1.533] lie entirely inside the observed range of structural holes, [0.789, 2.000]. The Lind–Mehlum (2010) test confirms a genuine inverted U[70]: the fitted slope is positive and significant at the lower bound (0.480, $z = 2.66$, $p < 0.01$) and negative

and significant at the upper bound (-0.416 , $z = -3.00$, $p < 0.01$), so the null of a monotone relationship is rejected ($p < 0.01$).

Table 3. Inverted-U turning point and Lind–Mehlum test for the structural-holes main effect.

Statistic	Value
Estimated turning point (structural-holes scale)	1.438
95% confidence interval (Delta method)	[1.347, 1.529]
95% confidence interval (Fieller method)	[1.275, 1.533]
Observed range of structural holes	[0.789, 2.000]
Slope at lower bound (0.789)	+0.480 ($z = 2.66$, $p < 0.01$)
Slope at upper bound (2.000)	-0.416 ($z = -3.00$, $p < 0.01$)
Lind–Mehlum test (H_0 : monotone vs. inverted U)	Rejected ($p < 0.01$)

The turning point also remains inside the observed range once the moderators are conditioned on. Under low, medium and high degree centrality, the conditional turning point is 1.459, 1.308 and 1.188, respectively; under low, medium and high closeness centrality, it is 1.268, 1.241 and 1.229. In every case, the point estimate and its 95% confidence interval fall within [0.789, 2.000], and the turning point shifts leftward as either centrality rises—the curve peaks earlier and steepens, which is the geometric counterpart of the moderations stated in H2a and H2b: an earlier, sharper downturn rather than a uniform downward shift in the curve.

5.3. Robustness Tests

Three alternative estimators are used to assess robustness, and results are reported in Table 4. First, negative binomial regression with robust (sandwich) standard errors is applied to guard against misspecification of the conditional variance (Models 1–3). Second, Poisson regression is applied as a standard count-data benchmark [61] (Models 4–6). Third, zero-inflated Poisson regression is applied to jointly address the count structure and the excess zeros (Models 7–9).

Across all three alternative estimators, the coefficients on structural holes (KSH) and its squared term (KSH^2) remain jointly significant with the signs predicted by H1. Specifically, Model 1 ($\beta = 1.064$, $p < 0.05$; $\beta = -0.128$, $p < 0.05$), Model 4 ($\beta = 1.097$, $p < 0.001$; $\beta = -0.131$, $p < 0.001$) and Model 7 ($\beta = 1.103$, $p < 0.001$; $\beta = -0.132$, $p < 0.001$) all confirm the inverted U-shaped relationship between knowledge-layer structural holes and breakthrough innovation [66].

The two steepening, peak-shifting moderations are similarly robust. The interaction between degree centrality and the squared structural-hole term ($KSH^2 \times CDC$) is significantly negative in Model 2 ($\beta = -0.138$, $p < 0.05$), Model 5 ($\beta = -0.102$, $p < 0.05$) and Model 8 ($\beta = -0.109$, $p < 0.05$), and the interaction between closeness centrality and the squared structural-hole term ($KSH^2 \times CCC$) is significantly negative in Model 3 ($\beta = -0.243$, $p < 0.001$), Model 6 ($\beta = -0.273$, $p < 0.001$) and Model 9 ($\beta = -0.288$, $p < 0.001$) [68,69]. The cross-layer integration mechanism and the optimal integration zone it delineates are therefore not artifacts of a particular estimator.

Table 4. Robustness test.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Methods of Analysis	Negative Binomial (Robust SE)			Poisson Regression		Zero-Inflated Poisson Regression			
Structural holes (KSH)	1.064 *	1.228 *	1.632	1.097 ***	1.251 ***	1.592 **	1.103 ***	1.231 ***	1.625 **
	(2.38)	(2.42)	(1.95)	(3.35)	(3.49)	(3.01)	(3.34)	(3.40)	(3.05)
Structural holes ² (KSH ²)	−0.128 *	−0.154 **	−0.215 *	−0.131 ***	−0.154 ***	−0.212 ***	−0.132 ***	−0.152 ***	−0.217 ***
	(−2.44)	(−2.58)	(−2.29)	(−3.43)	(−3.70)	(−3.56)	(−3.42)	(−3.62)	(−3.62)
Degree centrality (CDC)	0.151 ***	−1.260	0.140 ***	0.134 ***	−0.886	0.122 ***	0.136 ***	−0.987	0.124 ***
	(8.17)	(−1.82)	(7.65)	(12.30)	(−1.59)	(11.08)	(12.53)	(−1.74)	(11.35)
Closeness centrality (CCC)	−0.015	−0.000	−2.847 **	0.023	0.035	−3.249 ***	0.023	0.035	−3.422 ***
	(−0.36)	(−0.00)	(−2.72)	(0.86)	(1.26)	(−4.52)	(0.84)	(1.26)	(−4.69)
KSH × CDC		0.889 *			0.651 *			0.706 *	
		(2.18)			(2.01)			(2.15)	
KSH ² × CDC		−0.138 *			−0.102 *			−0.109 *	
		(−2.28)			(−2.16)			(−2.28)	
KSH × CCC			1.687 **			1.917 ***			2.020 ***
			(3.10)			(5.13)			(5.32)
KSH ² × CCC			−0.243 ***			−0.273 ***			−0.288 ***
			(−3.52)			(−5.65)			(−5.86)
Knowledge degree centrality (KDC)	−0.055 *	−0.041	−0.007	−0.057 **	−0.047 *	−0.006	−0.053 **	−0.042 *	0.003
	(−1.98)	(−1.45)	(−0.23)	(−2.92)	(−2.34)	(−0.27)	(−2.68)	(−2.05)	(0.14)
Knowledge closeness centrality (KCC)	−0.026	−0.004	−0.021	−0.029	−0.011	−0.023	−0.028	−0.010	−0.021
	(−0.72)	(−0.11)	(−0.55)	(−1.01)	(−0.38)	(−0.77)	(−0.98)	(−0.33)	(−0.73)
Betweenness centrality (CBC)	0.112 ***	0.112 ***	0.104 ***	0.094 ***	0.093 ***	0.087 ***	0.096 ***	0.096 ***	0.090 ***
	(5.96)	(6.05)	(6.15)	(16.58)	(16.50)	(15.01)	(17.05)	(16.98)	(15.53)
Amount of knowledge in existing patents	0.119 ***	0.121 ***	0.110 ***	0.109 ***	0.110 ***	0.105 ***	0.104 ***	0.104 ***	0.099 ***
	(3.69)	(3.76)	(3.59)	(6.50)	(6.53)	(6.24)	(6.16)	(6.19)	(5.82)
Patent stock	−0.055	−0.057 *	−0.052 *	−0.054 ***	−0.055 ***	−0.054 ***	−0.051 ***	−0.051 ***	−0.049 ***
	(−1.90)	(−1.99)	(−1.97)	(−4.18)	(−4.22)	(−4.11)	(−3.90)	(−3.94)	(−3.77)
Years of research	0.224 ***	0.226 ***	0.222 ***	0.207 ***	0.208 ***	0.203 ***	0.212 ***	0.213 ***	0.209 ***
	(12.15)	(12.19)	(12.03)	(12.83)	(12.90)	(12.63)	(13.06)	(13.13)	(12.87)
Constant	−3.468 ***	−3.780 ***	−4.463 *	−3.540 ***	−3.840 ***	−4.354 ***	−3.545 ***	−3.788 ***	−4.408 ***
	(−4.01)	(−3.80)	(−2.57)	(−5.50)	(−5.37)	(−3.95)	(−5.47)	(−5.25)	(−3.97)
Time-window fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5631	5631	5631	5631	5631	5631	5631	5631	5631
Log likelihood	−6038.60	−6034.58	−6022.63	−6138.22	−6134.84	−6113.48	−6131.64	−6128.20	−6104.77

Note:: z-statistics are in parentheses; all continuous predictors are standardized; * means $p < 0.05$, ** means $p < 0.01$, *** means $p < 0.001$. Additional sensitivity analyses. To further assess whether the results are driven by the operationalization of breakthrough innovation or by field-year heterogeneity in citation patterns, two additional sensitivity analyses are conducted. Sensitivity A applies a stricter top-1% breakthrough threshold and estimates a logit specification using the same control set. Sensitivity B re-estimates the models after standardizing the breakthrough measure by technology field and year, using negative binomial regression with year and technology-field fixed effects. The results are reported in Table 5.

In Sensitivity A, the inverted-U main effect remains evident under the stricter top-1% breakthrough proxy. The coefficient of knowledge-network structural holes is positive (KSH: $\beta = 1.784$, $z = 1.94$), while the squared term is negative and statistically significant (KSH²: $\beta = -0.220$, $p < 0.05$). More importantly, the estimated turning point on the original KSH scale is 1.405, and the Delta-method 95% confidence interval [1.267, 1.543] lies within the observed KSH range [0.789, 2.000]. The Lind–Mehlum test further supports the inverted-U interpretation: the slope is positive at the lower bound and negative at the upper bound, with both one-sided tests statistically significant. Thus, H1 is supported under the stricter breakthrough threshold, although the wider Fieller interval indicates that this tail-based specification should be interpreted as a demanding sensitivity check rather than as the primary curvature test.

The moderation results under Sensitivity A are also broadly consistent with the theoretical predictions. For degree centrality in the collaboration network, the interaction

between KSH² and CDC is negative and significant ($\beta = -0.438$, $p < 0.01$), indicating that higher collaboration-layer degree centrality steepens the inverted-U relationship and shifts its peak toward lower structural-hole levels. The conditional turning points under low, medium and high CDC are 1.45, 1.22 and 1.16, respectively, all falling within the observed KSH range. This supports H2a. For closeness centrality, the interaction between KSH² and CCC is also negative and significant ($\beta = -0.288$, $p < 0.05$), supporting H2b. The conditional turning points under medium and high CCC are both inside the observed range, while the confidence interval for the low-CCC condition is wider and not fully contained within the observed range. Therefore, the top-1% logit analysis supports the predicted negative closeness-centrality moderation, but the conditional curvature is less precisely estimated at the low-CCC level.

Sensitivity B provides a second check by addressing possible technology-field and year-specific citation heterogeneity. The results remain highly consistent with the main findings. The coefficient of KSH is positive and significant ($\beta = 1.012$, $p < 0.05$), and the coefficient of KSH² is negative and significant ($\beta = -0.121$, $p < 0.05$). The estimated turning point is 1.444 on the original KSH scale, with a Delta-method 95% confidence interval of [1.333, 1.555] and a Fieller 95% confidence interval of [1.097, 1.626], both fully inside the observed range [0.789, 2.000]. The Lind–Mehlum test again confirms that the curve rises at the lower bound and falls at the upper bound, supporting H1.

The two moderation effects are also robust in the field-year standardized negative binomial specification. The interaction between KSH² and CDC is negative and significant ($\beta = -0.142$, $p < 0.05$), supporting H2a. The conditional turning points under low, medium and high CDC are 1.48, 1.28 and 1.16, respectively, and all corresponding confidence intervals lie within the observed KSH range. Similarly, the interaction between KSH² and CCC is negative and highly significant ($\beta = -0.245$, $p < 0.001$), supporting H2b. The conditional turning points under low, medium and high CCC are 1.50, 1.23 and 1.22, respectively, again all falling within the observed range.

Overall, these two sensitivity analyses reinforce the main conclusion. The inverted-U relationship between knowledge-network structural holes and breakthrough innovation is preserved under both a stricter top-1% breakthrough threshold and a field-year standardized count specification. The degree-centrality moderation remains consistently negative across both checks. The closeness-centrality moderation is also negative in both analyses and statistically significant, although its conditional turning-point estimate is less precise under the low-closeness condition in the top-1% logit specification. These results indicate that the cross-layer integration mechanism is not an artifact of the baseline top-5% breakthrough definition, nor is it driven by field-year citation heterogeneity.

Table 5. Sensitivity analyses: stricter breakthrough threshold and field-year standardized specification.

Variable	Top-1% Logit			Field-Year Standardized NB		
	H1	H2a	H2b	H1	H2a	H2b
Structural holes (KSH)	1.784 (1.94)	2.236 * (2.16)	1.917 (1.38)	1.012 * (2.15)	1.181 * (2.21)	1.589 (1.83)
Structural holes ² (KSH ²)	−0.220 * (−2.04)	−0.294 * (−2.43)	−0.261 (−1.66)	−0.121 * (−2.19)	−0.148 * (−2.37)	−0.209 * (−2.16)
Degree centrality (CDC)	0.272 *** (5.77)	−4.328 * (−2.49)	0.270 *** (5.68)	0.152 *** (8.03)	−1.279 (−1.83)	0.141 *** (7.49)
Closeness centrality (CCC)	−0.024 (−0.32)	0.015 (0.19)	−3.635 (−1.91)	−0.016 (−0.38)	−0.002 (−0.04)	−2.836 ** (−2.63)
KSH × CDC		2.864 ** (2.80)			0.908 * (2.20)	
KSH ² × CDC		−0.438 ** (−2.92)			−0.142 * (−2.32)	
KSH × CCC			2.068 * (2.07)			1.689 ** (3.00)
KSH ² × CCC			−0.288 * (−2.23)			−0.245 *** (−3.42)
Amount of knowledge in existing patents	0.377 *** (4.15)	0.388 *** (4.26)	0.370 *** (4.10)	0.119 *** (3.70)	0.120 *** (3.76)	0.109 *** (3.57)

Patent stock	−0.253 *** (−3.40)	−0.266 *** (−3.60)	−0.252 *** (−3.42)	−0.056 (−1.93)	−0.058 * (−2.03)	−0.053 * (−1.99)
Years of research	0.607 *** (12.09)	0.610 *** (12.12)	0.605 *** (12.05)	0.223 *** (12.06)	0.225 *** (12.11)	0.220 *** (11.94)
Knowledge degree centrality (KDC)	−0.030 (−0.55)	0.002 (0.04)	0.020 (0.34)	−0.067 * (−2.38)	−0.054 (−1.87)	−0.019 (−0.61)
Knowledge closeness centrality (KCC)	−0.091 (−1.12)	−0.031 (−0.38)	−0.083 (−1.01)	−0.019 (−0.52)	0.007 (0.19)	−0.008 (−0.20)
Betweenness centrality in collaboration network (CBC)	0.421 *** (4.87)	0.405 *** (4.76)	0.380 *** (4.45)	0.112 *** (6.03)	0.112 *** (6.11)	0.104 *** (6.22)
Constant	−6.637 *** (−3.67)	−7.470 *** (−3.63)	−6.693 * (−2.34)	−3.398 *** (−3.82)	−3.742 *** (−3.67)	−4.453 * (−2.52)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Technology-field fixed effects	—	—	—	Yes	Yes	Yes
Observations	5631	5631	5631	5631	5631	5631
Turning point, original KSH scale	1.405			1.444		
Delta-method 95% CI	[1.267, 1.543]			[1.333, 1.555]		
Observed KSH range	[0.789, 2.000]			[0.789, 2.000]		
Lind–Mehlum test	Passed			Passed		

Note: z-statistics are reported in parentheses. The Top-1% Logit columns (H1, H2a, H2b) use a logit specification for the stricter top-1% breakthrough proxy; the Field-year standardized NB columns (H1, H2a, H2b) use negative binomial regression after field-year standardization and include both year and technology-field fixed effects. KSH denotes knowledge-network structural holes; CDC and CCC denote degree and closeness centrality in the collaboration network; KDC and KCC denote degree and closeness centrality in the knowledge network; CBC denotes betweenness centrality in the collaboration network. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

6. Discussion

Using multilayer network theory and panel data from the new energy field, this paper responded to the convergent challenges of declining global R&D productivity, the decisive decade for clean-energy transition, and the rapid restructuring of international innovation networks by theorizing and empirically validating a cross-layer integration mechanism. The three hypotheses were all supported: structural holes exhibit an inverted U-shaped relationship with breakthrough innovation (H1), and this relationship is negatively moderated by both degree centrality (H2a) and closeness centrality (H2b) in the collaboration network. Robustness tests using negative binomial regression with robust standard errors, Poisson and zero-inflated Poisson specifications produced consistent results, and a Lind–Mehlum test confirmed that the inverted-U turning point lies within the observed range of structural holes; the findings are also robust to a stricter top-1% breakthrough threshold.

6.1. Theoretical Extensions and Implications

The cross-layer integration mechanism advances structural hole theory and multi-layer network theory along three connected dimensions, each of which can be sharpened by direct contrast with the prior literature.

First, structural holes are repositioned from a within-layer information advantage to a cross-layer entry channel. The dominant tradition following Burt (1992) and Ahuja (2000) treats structural holes as a static positional benefit: nodes that span them obtain more heterogeneous information and, by linear extension, more innovation [7,22]. Even studies that introduce nonlinearity, such as Kaplan and Vakili (2015) on recombinant breakthroughs [36,71], locate the diminishing returns within a single network. The present study departs from both lines by showing that heterogeneous information is not valuable on its own; it becomes valuable only once it crosses into the collaboration layer where

recombination actually occurs. The inverted U-shaped curve is then no longer a within-layer artifact but the natural consequence of a cross-layer transfer disciplined by a finite cognitive processing threshold—a result that single-layer designs are structurally incapable of producing.

Second, collaboration-layer degree and closeness centrality are reframed from access indicators to gates of cross-layer receptivity. In the single-layer accounts of Freeman, Ibarra and subsequent work, these centralities are typically associated with faster information flow and are, at the margin, presumed to support innovation [23,24]. Single-layer studies that report negative effects of high centrality (e.g., Powell et al., 1996 on biotech [13]) attribute them to local saturation, but stop short of specifying the structural channel through which the suppression operates. The present study identifies that channel: when either centrality becomes excessive, the collaboration layer drifts into homogenization (high degree centrality) or localization (high closeness centrality) and loses the capacity to receive, distribute and recombine heterogeneous information arriving from the knowledge layer. The resulting moderation is therefore not a generic “weakening”: it specifically compresses the inverted U-shaped main effect—shifting its peak toward lower structural-hole levels and steepening the post-peak decline—which is exactly the pattern observed in Models 3 and 4 and in Figures 7 and 8.

Third, the joint operation of the decaying main effect and the two steepening, peak-shifting moderations delineates an optimal cross-layer integration zone—a construct that single-layer analyses cannot generate. Existing multilayer studies in the Boccaletti and Kivela tradition [8,9] have established rich formal and topological properties of multilayer networks but stop at structural description, while management-side multilayer work tends to verify “layers matter” without specifying how. Our contribution is to convert “layers matter” into a testable mechanism: one layer supplies heterogeneous information, the other layer must receive and recombine it, and the two boundary forces—cognitive threshold within the inventor and structural diversity within the receiving layer—jointly define a region in parameter space within which breakthrough innovation is maximized. Outside this zone, breakthrough innovation is suppressed through one of two distinct channels: either the main-effect curve decays (too few or too many structural holes), or the curve steepens and its peak shifts to a lower structural-hole level, narrowing the productive window (excessive collaboration-layer centrality). By bringing this characterization to a fixed-effects negative binomial specification on 5631 inventor-year observations and three robustness estimators, the paper also moves multilayer network research from structural description toward mechanism-level identification, providing a template for future mechanism-oriented multilayer analyses.

6.2. Alternative Explanations, Endogeneity, and the Scope of the Findings

The inverted U-shape is interpreted above through inventors’ finite cognitive processing capacity, but this is not the only mechanism consistent with the curvature. As the heterogeneity of the knowledge an inventor must integrate grows, the coordination costs of joint work, the managerial complexity of running broader projects, and the organizational constraints on combining increasingly distant technological domains all rise as well, and any of these could in principle generate the same pattern of diminishing and then negative returns. The present design cannot empirically separate cognitive overload from these coordination- and organization-based accounts, because patent-panel data record the structural traces of integration rather than the cognitive or organizational processes behind them. We therefore treat the cognitive-processing threshold as one theoretically grounded mechanism that is consistent with the evidence, not as the uniquely identified cause, and we read the two collaboration-layer moderations in the same spirit—as

evidence on how the structure of the receiving layer conditions integration, rather than as a direct measurement of receptivity.

A related caution concerns endogeneity. The five-year lag between the network measures and the breakthrough outcomes, together with the fixed-effects specification, mitigates simple reverse causality and time-invariant confounding, but it does not remove the possibility of dynamic feedback. Inventors who have already produced exceptional breakthroughs may subsequently attract more collaborators, gain access to broader knowledge, and come to occupy more advantageous structural-hole positions, so that network position and innovation reinforce each other over time. Because the design contains no exogenous shock or instrument for network position, the estimates are best read as identifying the shape and boundary conditions of the cross-layer mechanism rather than as fully exogenous causal effects. What the evidence supports is that, conditional on observed covariates and time-invariant heterogeneity, the relationship between knowledge-layer structural holes and breakthrough innovation is curvilinear and is systematically reshaped by collaboration-layer centrality; what it cannot establish is the precise counterfactual effect of an externally imposed change in an inventor's network position. Disentangling these effects—ideally with quasi-experimental variation, or with the micro-level data discussed in Section 6.4—is an important task for future research.

6.3. Management Recommendations

The cross-layer integration mechanism offers several concrete recommendations for managers of Technology Innovation Composite Networks.

First, managers should maintain structural diversity in the collaboration layer and avoid over-reliance on a few high-centrality nodes. While high centrality facilitates rapid information flow, excessive concentration homogenizes the collaboration environment and erodes its capacity to receive heterogeneous information from the knowledge layer. A balanced collaboration network, built on multi-point cooperation among inventors, preserves the receptive capacity needed for structural holes to translate into breakthrough outputs [40,41].

Second, managers should actively counter localization by promoting cross-departmental and cross-organizational collaboration through joint R&D projects and cross-training. Excessively high closeness centrality confines information flow to immediate neighbors and prevents heterogeneous information from diffusing across distant parts of the collaboration network, thereby cutting short the cross-layer pathway on which breakthrough innovation depends [24].

Third, managers should institutionalize open communication platforms—regular cross-departmental meetings and innovation information-sharing systems—to reduce the dual risks of homogenization and localization. Such platforms broaden inventors' collaborator bases and help ensure that heterogeneous knowledge surfaced by structural-hole spanning is circulated widely rather than absorbed into a narrow community [40,41].

Finally, managers should treat the network structure as dynamically managed rather than as a static property. Structural holes, degree centrality and closeness centrality co-determine where the composite network lies relative to the optimal cross-layer integration zone, and that position shifts as the network grows and inventors enter and exit. Periodic reviews of the network structure, with targeted interventions when the collaboration layer drifts toward centralization or localization, keep the system closer to the peak of the inverted U-shaped curve and thus closer to the conditions that sustain breakthrough innovation [41,72].

Beyond firm managers, these prescriptions also speak to policymakers and to the coordinators of large-scale R&D platforms. Policymakers designing clean-energy innovation programs can promote controlled knowledge diversity by funding cross-domain and

cross-organizational consortia rather than concentrating support on a few dominant incumbents, and by rewarding genuinely novel recombination rather than patent counts alone, so that the composite network does not drift toward the over-centralized and over-localized configurations that compress the structural-hole advantage. Coordinators of shared R&D platforms can operationalize the optimal cross-layer integration zone more directly: by rotating bridging and broker roles instead of letting a few hub inventors monopolize them, by setting up cross-team matching and shared knowledge repositories that move heterogeneous information out of local cliques, and by monitoring layer-level indicators—degree and closeness centralization in the collaboration layer, and the supply of structural holes in the knowledge layer—so that targeted interventions can be triggered when either layer drifts away from the productive region. Because high collaboration-layer centrality narrows the band of structural holes that can be productively absorbed, all three audiences should pair any push for more heterogeneous knowledge with deliberate measures to keep the collaboration layer diverse and globally connected.

6.4. Limitations and Future Research

First, the empirical setting is confined to the new energy field over 2004–2018. New energy is a useful testbed because more than 64.80% of its patents integrate multiple knowledge elements and exhibit rich collaboration relationships [47], which makes the cross-layer integration mechanism observable, but the same density of cross-domain recombination may not characterize industries with shorter innovation cycles (e.g., consumer electronics) or industries with more concentrated knowledge bases (e.g., pharmaceuticals). Future work should re-examine the inverted U-shape and the two steepening, peak-shifting moderations across industries with different innovation cycles and knowledge-base structures, to delineate the boundary conditions under which the optimal cross-layer integration zone preserves its shape.

Second, the cognitive processing threshold is identified at the network level rather than at the individual level. Our design infers the threshold from the curvature of the inverted U, but does not directly measure how individual inventors absorb, evaluate and recombine heterogeneous information. This is an unavoidable consequence of patent-based panel data, in which cognitive processes are observed only through their structural traces. Future research could combine the network-level evidence developed here with micro-level data—laboratory experiments, inventor interviews, or fine-grained R&D activity logs—to directly recover the threshold and to test whether it varies systematically with inventor characteristics such as experience or disciplinary breadth.

Third, the TICN treats the knowledge layer and the collaboration layer as the only two layers in the system, abstracting away from the rapid emergence of AI systems as active participants in research and invention. Foundation models, autonomous experimentation systems and AI-augmented ideation tools are increasingly involved in literature synthesis, hypothesis generation and experimental design, and their presence is beginning to reshape both layers studied here. The dual-layer TICN generalizes naturally to a tri-layer architecture in which AI agents form a third, distinct layer with its own connectivity structure and informational reach. Within this extension, the cross-layer integration mechanism developed here—heterogeneous information supplied by one layer, received and recombined by another, bounded upward by a cognitive processing threshold and outward by the receiving layer's structural features—provides a ready-made analytical scaffold for two timely questions. First, does the inventor's cognitive processing threshold expand when AI agents absorb part of the heterogeneity-screening burden, thereby shifting the peak of the inverted U-shaped curve rightward? Second, do the window-narrowing effects of collaboration-layer centralities attenuate when AI agents bypass localized human collaboration cliques and reconnect otherwise isolated parts of the network?

Answering these questions would extend the optimal cross-layer integration zone from a two-layer characterization to a three-layer one and would place the present mechanism at the center of the post-LLM R&D paradigm.

7. Conclusions

This study sets out to specify how knowledge-layer structural holes yield breakthrough innovation, and under what conditions. Treating structural holes as cross-layer entry channels and collaboration-layer centralities as gates of receptivity, it showed—on 5631 inventor-year observations from the new energy field and across four estimators—that the structural-hole effect on breakthrough innovation is inverted U-shaped, and that degree and closeness centrality steepen this curve and shift its peak toward lower structural-hole levels, narrowing the window over which cross-layer integration remains productive. These results move multilayer innovation research from establishing that layers matter to specifying how information must cross between them, and they imply that managers, policymakers and platform coordinators should match the intensity of structural holes to the state of the collaboration layer rather than maximize either in isolation. The mechanism is identified at the network level and under the design’s lag-and-fixed-effects assumptions rather than experimentally; extending it with micro-level evidence, with quasi-experimental variation, and to the emerging three-layer architecture in which AI agents form a distinct layer is the natural next step.

Author Contributions: Conceptualization, S.L.; Methodology, S.L., J.D. and Z.W.; Software, Z.W.; Validation, J.D. and M.S.; Formal analysis, Z.W.; Investigation, Z.Y.; Resources, X.Y.; Data curation, Z.Y.; Writing—original draft, S.L. and Z.W.; Writing—review & editing, J.D., M.S. and X.Y.; Visualization, Z.Y.; Supervision, S.L., M.S. and X.Y.; Project administration, M.S. and X.Y.; Funding acquisition, S.L.. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of this study are not publicly available because the underlying patent records were obtained from the Derwent Patent Database under license and cannot be redistributed. The search strategy, inclusion criteria, inventor-disambiguation rules, IPC-coding scheme and matrix-construction procedures are described in Section 4.1 to support reproducibility.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

TICN	Technology Innovation Composite Network
SHT	Structural Hole Theory
ZINB	Zero-Inflated Negative Binomial
IPC	International Patent Classification
AI	Artificial Intelligence
R&D	Research and Development
EU	European Union
SME	Small and Medium-sized Enterprise
NB	Negative Binomial
KSH	Knowledge-network Structural Holes

CDC	Collaboration-network Degree Centrality
CCC	Collaboration-network Closeness Centrality
CBC	Collaboration-network Betweenness Centrality
KDC	Knowledge-network Degree Centrality
KCC	Knowledge-network Closeness Centrality

References

- Li, S.; Wen, Z. Interaction Effect of Knowledge Network and Collaboration Network on Breakthrough Innovation. In *Proceedings of 5th International Communication Engineering and Cloud Computing Conference, CECCC 2023, Nanjing, China, 27–29 October 2023*; Institute of Electrical and Electronics Engineers Inc.: Nanjing, China, 2023.
- European Commission. *Horizon 2020: The EU Framework Programme for Research and Innovation*; European Commission: Brussels, Belgium, 2020.
- National Science & Technology Council. *National AI R&D Strategic Plan*; National Science & Technology Council: Washington, DC, USA, 2019.
- Freeman, C. *Technology Policy and Economic Performance: Lessons from Japan*; Pinter Publishers: London, UK, 1987.
- Schumpeter, J.A. *Business Cycles: A Theoretical, Historical, and Statistical Analysis of the Capitalist Process*; McGraw-Hill: New York, NY, USA, 1939.
- Granovetter, M. Economic action and social structure: The problem of embeddedness. *Am. J. Sociol.* **1985**, *91*, 481–510.
- Burt, R.S. *Structural Holes: The Social Structure of Competition*; Harvard University Press: Cambridge, MA, USA, 1992.
- Boccaletti, S.; Bianconi, G.; Criado, R.; del Genio, C.I.; Gómez-Gardeñes, J.; Romance, M.; Sendiña-Nadal, I.; Wang, Z.; Zanin, M. The structure and dynamics of multilayer networks. *Phys. Rep.* **2014**, *544*, 1–122.
- Kivelä, M.; Arenas, A.; Barthélemy, M.; Gleeson, J.P.; Moreno, Y.; Porter, M.A. Multilayer networks. *J. Complex Netw.* **2014**, *2*, 203–271.
- Powell, W.W.; Grodal, S. Networks of innovators. In *The Oxford Handbook of Innovation*; Oxford University Press: Oxford, UK, 2005; pp. 56–85.
- Schilling, M.A.; Phelps, C.C. Interfirm collaboration networks: The impact of large-scale network structure on firm innovation. *Manag. Sci.* **2007**, *53*, 1113–1126.
- Schilling, M.A. *Strategic Management of Technological Innovation*; McGraw-Hill Education: Columbus, OH, USA, 2010.
- Powell, W.W.; Koput, K.W.; SmithDoerr, L. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Adm. Sci. Q.* **1996**, *41*, 116–145.
- Granovetter, M.S. The strength of weak ties. *Am. J. Sociol.* **1973**, *78*, 1360–1380.
- Jin, X.; Yu, Y.; Zhang, M.; Chen, C.; Li, Y. Two-Tier Network Embeddedness, Heterogeneous Resource Acquisition, and Firms' Breakthrough Innovation: The Moderating Effect of Digitalization. *Systems* **2025**, *13*, 1012.
- Song, L.; Chen, S.; Liang, J.; Yin, X. Knowledge Recombination Reveals the Nonlinear Influence of Team Scale on Technological Breakthroughs. *Systems* **2025**, *13*, 877.
- Ding, L.; Lawson, C.; Shapira, P. Rise of Generative Artificial Intelligence in Science. *Scientometrics* **2025**, *130*, 5093–5114.
- Boussiou, L.; Lane, J.N.; Zhang, M.; Jacimovic, V.; Lakhani, K.R. The Crowdless Future? Generative AI and Creative Problem-Solving. *Organ. Sci.* **2024**, *35*, 1589–1607.
- Fleming, L. Recombinant uncertainty in technological search. *Manag. Sci.* **2001**, *47*, 117–132.
- Henderson, R.M.; Clark, K.B. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Adm. Sci. Q.* **1990**, *35*, 9–30.
- Dahlin, K.B.; Behrens, D.M. When is an invention really radical? Defining and measuring technological radicalness. *Res. Policy* **2005**, *34*, 717–737.
- Ahuja, G. Collaboration networks, structural holes, and innovation: A longitudinal study. *Adm. Sci. Q.* **2000**, *45*, 425–455.
- Freeman, L.C. *Centrality in Social Networks: Conceptual Clarification*, in *Social Networks: Critical Concepts in Sociology*; Routledge: Oxfordshire, UK, 2002; Volume 1, pp. 238–263.
- Ibarra, H. Network centrality, power, and innovation involvement: Determinants of technical and administrative roles. *Acad. Manag. J.* **1993**, *36*, 471–501.
- Hall, B.H.; Jaffe, A.B.; Trajtenberg, M. The NBER patent citations data file: Lessons, insights, and methodological tools. In *NBER Working Paper Series*; National Bureau of Economic Research: Cambridge, MA, USA, 2001.
- Griliches, Z. Patent statistics as economic indicators: A survey. *J. Econ. Lit.* **1990**, *28*, 1661–1707.

27. Yang, H.; An, S. Critical nodes identification in complex networks. *Symmetry* **2020**, *12*, 123.
28. Jialu, S.; Zhiqiang, M.; Binxin, Z.; Haoyang, X.; Oteng, A.F.; Hu, W. Collaborative Innovation Network, Knowledge Base, and Technological Innovation Performance-Thinking in Response to COVID-19. *Front. Psychol.* **2021**, *12*, 648276.
29. Lee, A.R.; Son, S.M.; Kim, K.K. Information and communication technology overload and social networking service fatigue: A stress perspective. *Comput. Hum. Behav.* **2016**, *55*, 51–61.
30. Ouyang, Y.; Ye, F.; Tan, K. The effect of strategic synergy between local and neighborhood environmental regulations on green innovation efficiency: The perspective of industrial transfer. *J. Clean. Prod.* **2022**, *380*, 134933.
31. Abhari, K.; Davidson, E.J.; Xiao, B. Collaborative innovation in the sharing economy: Profiling social product development actors through classification modeling. *Internet Res.* **2019**, *29*, 1014–1039.
32. Wang, Y.; Li, N.; Zhang, B.; Huang, Q.; Wu, J.; Wang, Y. The effect of structural holes on producing novel and disruptive research in physics. *Scientometrics* **2023**, *128*, 1801–1823.
33. Xu, L.; Ling, J.; Park, H.D. Founding team homogeneity and network positioning: The moderating role of environmental and organizational factors. *J. Small Bus. Manag.* **2023**, *61*, 1773–1805.
34. Guo, M.; Yang, N.; Wang, J. How do structural holes promote network expansion? *Technol. Forecast. Soc. Chang.* **2021**, *173*, 121129.
35. Shipilov, A.V.; Li, S.X. Can you have your cake and eat it too? Structural holes' influence on status accumulation and market performance in collaborative networks. *Adm. Sci. Q.* **2008**, *53*, 73–108.
36. Kaplan, S.; Vakili, K. The double-edged sword of recombination in breakthrough innovation. *Strateg. Manag. J.* **2015**, *36*, 1435–1457.
37. Li, S.G.; Zhu, H. Co-creation of metamere brands equity with the impact of celebrity virtual images and real-life influence: From the perspective of experience-dominant logic. *Electron. Commer. Res.* **2024**, *26*, 213–242.
38. Badar, K.; Hite, J.M.; Badir, Y.F. The moderating roles of academic age and institutional sector on the relationship between co-authorship network centrality and academic research performance. *Aslib J. Inf. Manag.* **2014**, *66*, 38–53.
39. Min, C.; Bu, Y.; Sun, J.J. Predicting scientific breakthroughs based on knowledge structure variations. *Technol. Forecast. Soc. Chang.* **2021**, *164*, 120502.
40. Wu, W.L.; Yeh, R.S.; Hung, H.K. Knowledge Sharing and Work Performance: A Network Perspective. *Soc. Behav. Personal.* **2012**, *40*, 1113–1120.
41. Huang, Y.; Hong, T.; Ma, T. Urban network externalities, agglomeration economies and urban economic growth. *Cities* **2020**, *107*, 102882.
42. Smojver, V.; Štorga, M.; Zovak, G. Exploring knowledge flow within a technology domain by conducting a dynamic analysis of a patent co-citation network. *J. Knowl. Manag.* **2021**, *25*, 433–453.
43. Hwang, S.; Shin, J. Extending technological trajectories to latest technological changes by overcoming time lags. *Technol. Forecast. Soc. Chang.* **2019**, *143*, 142–153.
44. Yao, L. New energy utilization in environmental design and realization. *Energy Rep.* **2022**, *8*, 9211–9220.
45. Zhang, J.; Huang, W.L. A Pilot Assessment of New Energy Usage Behaviors: The Impacts of Environmental Accident, Cognitions, and New Energy Policies. *Front. Environ. Sci.* **2022**, *10*, 1035.
46. Cheng, L.; Liu, Y.; Lou, X. Does technology conglomeration promote innovative outcomes of new energy vehicle enterprises? The moderating effect of divisive faultlines. *J. Clean. Prod.* **2021**, *324*, 129232.
47. Yayavaram, S.; Ahuja, G. Decomposability in knowledge structures and its impact on the usefulness of inventions and knowledge-base malleability. *Adm. Sci. Q.* **2008**, *53*, 333–362.
48. Capponi, G.; Martinelli, A.; Nuvolari, A. Breakthrough innovations and where to find them. *Res. Policy* **2022**, *51*, 104376.
49. Randle, D.K.; Pisano, G.P. The Evolutionary Nature of Breakthrough Innovation: An Empirical Investigation of Firm Search Strategies. *Strategy Sci.* **2021**, *6*, 290–304.
50. Jaffe, A.B.; De Rassenfosse, G. *Patent Citation Data in Social Science Research: Overview and Best Practices*; Edward Elgar Publishing: Cheltenham, UK, 2019.
51. Burt, R.S. Structural holes and good ideas. *Am. J. Sociol.* **2004**, *110*, 349–399.
52. Burt, R.S. Structural holes capstone, cautions, and enthusiasms. *Pers. Netw. Class. Read. New Dir. Egocentric Anal.* **2021**, *51*, 384.
53. Zhang, J.; Luo, Y. Degree centrality, betweenness centrality, and closeness centrality in social network. In Proceedings of MSAM2017, Bangkok, Thailand, 26–27 March 2017.
54. Evans, T.S.; Chen, B. Linking the network centrality measures closeness and degree. *Commun. Phys.* **2022**, *5*, 172.
55. Bedford, A.; Ma, L.; Ma, N. Patenting activity or innovative originality? *Account. Financ.* **2021**, *61*, 4191–4207.

56. Wang, X.X.; Jiao, H.Y. The impact of network positions in scientific collaboration on pharmaceutical firms' technological innovation performance: Moderating roles of scientific collaboration strength and patent stock. *Front. Public Health* **2022**, *10*, 980845.
57. Rungta, P.D.; Meena, C.; Sinha, S. Identifying nodal properties that are crucial for the dynamical robustness of multistable networks. *Phys. Rev. E* **2018**, *98*, 022314.
58. Srivastava, M.K.; Gnyawali, D.R.; Hatfield, D.E. Behavioral implications of absorptive capacity: The role of technological effort and technological capability in leveraging alliance network technological resources. *Technol. Forecast. Soc. Change* **2015**, *92*, 346–358.
59. Li, E.Y.; Liao, C.H.; Yen, H.R. Co-authorship networks and research impact: A social capital perspective. *Res. Policy* **2013**, *42*, 1515–1530.
60. Dong, J.Q.; Yang, C.H. Information technology and innovation outcomes: Is knowledge recombination the missing link? *Eur. J. Inf. Syst.* **2019**, *28*, 612–626.
61. Cameron, A.C.; Trivedi, P.K. *Regression Analysis of Count Data*; Cambridge University Press: Cambridge, UK, 2013.
62. Hausman, J.A.; Hall, B.H.; Griliches, Z. *Econometric Models for Count Data with an Application to the Patents-R&D Relationship*; Econometrica, Bethesda, MD, USA 1984.
63. Gulati, R.; Sytch, M.; Tatarynowicz, A. The rise and fall of small worlds: Exploring the dynamics of social structure. *Organ. Sci.* **2012**, *23*, 449–471.
64. Rose, C.E.; Martin, S.W.; Wannemuehler, K.A.; Plikaytis, B.D. On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data. *J. Biopharm. Stat.* **2006**, *16*, 463–481.
65. He, H.; Wang, W.; Hu, J. Distribution-free inference of zero-inflated binomial data for longitudinal studies. *J. Appl. Stat.* **2015**, *42*, 2203–2219.
66. Osgood, D.W.; Finken, L.L.; McMorris, B.J. Analyzing multiple-item measures of crime and deviance II: Tobit regression analysis of transformed scores. *J. Quant. Criminol.* **2002**, *18*, 319–347.
67. Jacobson, T.; Zou, H. High-dimensional censored regression via the penalized tobit likelihood. *J. Bus. Econ. Stat.* **2024**, *42*, 286–297.
68. Mu, J.; Di Benedetto, A. Networking capability and new product development. *IEEE Trans. Eng. Manag.* **2011**, *59*, 4–19.
69. Johnston, D.W.; Piatti, M.; Torgler, B. Citation success over time: Theory or empirics? *Scientometrics* **2013**, *95*, 1023–1029.
70. 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.
71. Soda, G.; Stea, D.; Pedersen, T. Network Structure, Collaborative Context, and Individual Creativity. *J. Manag.* **2019**, *45*, 1739–1765.
72. Goldstein, R.; Vitevitch, M.S. The Influence of Closeness Centrality on Lexical Processing. *Front. Psychol.* **2017**, *8*, 1683.

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.