Journal of Informetrics 19 (2025) 101611

Contents lists available at ScienceDirect

Journal of Informetrics

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

The effects of scientific collaboration network structures on impact and innovation: A perspective from project teams

Zhifeng Liu [a], Chenlin Wang [a,*], Jinqing Yang [b]

[a] *Department of Information Management, Peking University, Beijing, China*
[b] *School of Information Management, Central China Normal University, Wuhan, China*

ARTICLE INFO

Keywords:
Scientific collaboration
Project team
Structural hole
Small-world

ABSTRACT

Scientific collaboration is critical in tackling complex research challenges, necessitating optimized configurations of research teams. While existing research primarily examines the impact of collaboration network characteristics on the impact and innovation of individual papers, there is less focus on these characteristics within the context of research projects. To bridge this gap, this study adopts the perspective of project teams and explores the influence of scientific collaboration network structures on the impact and innovation of research outputs. By employing ordinary least squares regression and negative binomial regression methods on a dataset encompassing 21,618 NSF grants and their associated 351,550 publications, we rigorously analyze how specific network characteristics impact the innovation and impact of the research outputs. The results reveal a negative correlation between the count of structural holes and both the impact and conventionality of the team's papers. Meanwhile, the small world of a project team positively correlates with the papers' impact and displays an inverted U-shaped relationship with innovation. Further analysis confirms that there is no interactive effect between structural holes and small world. A series of robustness checks have been conducted, demonstrating that these findings are robust. This study contributes valuable insights for scholars, institutions, and policymakers aiming to enhance research team effectiveness. It underscores the nuanced impacts of network properties on research outputs, offering a new perspective by focusing on project-based team structures rather than individual paper collaborations.

1. Introduction

Research projects serve as the fundamental units of scientific research and the core vehicles for technological innovation activities (Yang, 2022), the impact and the innovation performance of which have garnered widespread attention. Scientific collaboration is a set of informal functions and formal activities among scientists engaged in knowledge production, particularly participating in research projects (De Stefano et al., 2013). Literature has indicated a close correlation between scholarly collaboration and research productivity, impact, and innovation performance (Vieira, 2023). Most research efforts have examined papers of individual collaborators as sources for understanding the structure of collaboration networks (Park et al., 2022). Scholars have explored the connection between structural characteristics of scientific collaboration networks and the impact or innovation performance of scientific outputs, examining angles like peer collaboration (Walter et al., 2022), cross-regional cooperation (Zhou & Li, 2023), and beyond. In contrast,

* Corresponding author.
E-mail address: wangcl@stu.pku.edu.cn (C. Wang).

https://doi.org/10.1016/j.joi.2024.101611
Received 28 January 2024; Received in revised form 6 September 2024; Accepted 13 November 2024
Available online 18 November 2024
1751-1577/© 2024 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

collaboration networks based on research projects encompass broader research fields and involve more researchers, thereby better capturing cross-disciplinary influences and knowledge flows (Zou et al., 2023). These collaboration networks of research projects embed rich scientific capital and social capital (Benz & Rossier, 2022), providing a more accurate reflection of the characteristics and mechanisms of team collaboration.

Common network structural features include network scale, node degree, centrality, clustering coefficient, average shortest path, structural holes, and small world. Among them, structural holes, serving as bridges connecting different subgroups (Burt, 1992), play a vital role in fostering knowledge fusion and innovation among diverse groups. The notion of small-world networks stands out for its efficiency and equity in disseminating knowledge (Kim & Park, 2009). In scientific collaboration networks formed based on research projects, the actors include researchers who are members of projects. Structural holes represent the information gap between unconnected project members within the network (Brunetta et al., 2015), and the more structural holes an actor bridges, the richer their information gains. Small-world phenomenon represent the close connections among project members, indicating unique information transmission capabilities, which is explained by the emergence of frequently publishing researchers with many distinct co-authors (Rose, 2022). Therefore, structural holes and small world emphasize the efficiency of information flow compared to other network features (Su & Jiang, 2023), which is crucial for enhancing the performance of research teams. While these two features affect information flow in different ways, collectively shaping the network structure and the flowing effect. Thus, there may be some relations between them, asking for further exploration. This study examines both features, as well as their interactions and combined effects, to better understand how project collaboration network structures influence the impact and innovation of team outputs.

Considering the above points, we formulate three research questions:

RQ1: How do structural holes within scientific project team network affect the impact and innovation of project outputs?

RQ2: How does small-world phenomenon of scientific project team network bear upon the impact and innovation of project outputs?

RQ3: What is the influence of the interaction between structural holes and small world on the impact and novelty of project outputs?

This study primarily contributes to the following three aspects. Firstly, by utilizing research projects as the foundational unit for examining research teams, this work offers a fresh perspective, departing from traditional studies that typically focus on individual research papers. Secondly, the research illuminates how structural holes and small world within collaboration networks of research project teams influence the impact and innovation of research outputs, affording a more profound comprehension of the intricacies and efficacy of collaboration networks. Thirdly, it provides targeted recommendations for optimizing the collaboration strategies of research project teams and enhancing the impact and innovation of research outputs.

2. Related research and hypotheses

2.1. The structural holes of project team collaboration networks and impact

The architecture of a scientist's collaborative network (Fronczak et al., 2022) is profoundly linked to their scholarly accomplishments. Structural hole is an important network structural feature, acting as bridges between different subgroups and representing opportunities for the flow of information, resources, and ideas between different subgroups (Burt, 1992). However, they might also become points of congestion. On the one hand, social friction is inherent in the architecture of structural holes, and interacting with subgroups simultaneously may require switching between different cognitive frameworks to address the conflict (Halevy et al., 2019), causing a decrease in the efficiency of information transmission. On the other hand, structural holes may encourage opportunistic behaviors among team members, hindering the flow of information (Xiang et al., 2023).

Within collaboration networks, nodes that bridge structural holes win a competitive edge due to access to non-redundant resources (Burt, 1992). Nevertheless, these boundary spanners also experience role conflicts. For team members within the collaboration network, the impact of their academic outputs is associated with their node attributes. The structural holes spanned by teams result from their historical status and centrality, along with the structural holes bridged in the past by team members (Zaheer & Soda, 2009). Researchers situated in more central positions or spanning a greater number of structural holes within a collaborative network generally achieve higher impact than their less centrally-located counterparts (Tahmooresnejad et al., 2021). However, some studies have presented contrasting findings. For instance, in the knowledge and collaboration network within the wind energy domain, structural holes exhibit a positive but statistically insignificant effect on output impact, while members' centrality demonstrates an inverted U-shaped relationship with output impact (Guan et al., 2017a).

Structural holes in the co-author network can hinder effective knowledge sharing and dissemination within the team. Within teams, structural holes of leaders are negatively associated with team performance (Li et al., 2017). Expert teams generate less impactful outputs when they occupy a greater number of structural holes (Schillebeeckx et al., 2019). Excessive structural holes can lead to information blockages and inadequate flow, restricting the dissemination and citation of papers across different subgroups, thus reducing the impact of team outputs.

Hypothesis 1. (H1): Structural holes in project team collaboration networks have a negative relationship with impact of outputs.

2.2. The structural holes of project team collaboration networks and innovation

For R&D project teams operating under constrained resources, seamless coordination, collaboration, and idea exchange are imperative to maintain innovation (Hung, 2017). Innovation is the process in which researchers creatively utilize existing knowledge to generate new knowledge or ideas resulting from recombining existing knowledge (Galunic & Rodan, 1998). Successful innovation requires a balance between conventionality and novelty (Dai et al., 2023). Conventionality reflects the degree to which research outcomes adhere to recognized conventional frameworks (O'Connor, 2021), while novelty is an essential feature of creative ideas. However, novelty and conventionality are not opposing factors in the production of science (Uzzi et al., 2013). Increasing conventionality helps bridge the information gap created by novelty, making new ideas more easily accepted. New ideas need to be embedded in a framework of accepted conventions, which increase the audience's ability to evaluate innovation in an ever-expanding sea of knowledge fluently (Mukherjee et al., 2016). Uzzi et al. (2013) developed a methodology based on paper citation networks and journal pairs to understand scientific innovation. They identified novel pairings through atypical knowledge combinations and conventional pairings through relatively common combinations. An indicator Z-score generated from journal pairs was used to quantify the conventionality and novelty of a paper, corresponding to positive and negative values, respectively. This method was employed to examine the innovation of millions of research articles in the Web of Science (WOS). Thus, it can also be used to measure the innovation of output papers within project teams.

Studies suggest that network attributes are pivotal in driving innovation (Yang et al., 2021). Even in situations where strong ties prevail, weak network structures like structural holes harness the power of strong ties for innovation creation (Rost, 2011). Structural holes, acting as conduits linking various subgroups in a network, foster interdisciplinary collaboration (Guan et al., 2017b). Those individuals spanning structural holes are not only privy to unique information but are also predisposed to conceive transformative and inventive ideas (Wang et al., 2023). Nonetheless, the impact is not universally positive. In R&D collaboration networks, structural holes and network prominence critically influence exploratory innovation, exhibiting an upside-down U-shaped relationship. Occupying the two positions simultaneously hampers exploratory innovation (Ma et al., 2020). This study posits that structural holes can facilitate the introduction of perspectives from diverse backgrounds into a team. This can catalyze fresh cognitive paradigms and groundbreaking research propositions. Furthermore, structural holes serve as catalysts for knowledge dissemination and collaborations across diverse subgroups, potentially culminating in the formulation of more groundbreaking publications. Thus, we posit the following hypothesis:

Hypothesis 2a. (H2a): Structural holes in project team collaboration networks have a positive relationship with novelty of outputs.

Hypothesis 2b. (H2b): Structural holes in project team collaboration networks have a negative relationship with conventionality of outputs.

2.3. The small world of project team collaboration networks and impact

Network cohesion is highlighted as the key to community strength (Moody, 2004). Two pivotal metrics in social network analysis are the clustering coefficient and the average shortest path, representing local cohesion and overall connectivity, respectively. A network characterized by high clustering and short path length is termed a small-world network (Watts & Strogatz, 1998). The small-world phenomenon emerges through network stars, authors serving as intermediaries among numerous coauthors (Andrikopoulos et al., 2020). Interlinked stars connect various parts of the network through their collaborations, which can access more information as bridges, thereby expanding their reach of knowledge and amplifying the impact of the collaboration network. Confirming this, Ebadi and Schiffauerova (2015) found a positive correlation between small-world networks and article quality, assessed by citation counts and journal impact factors.

Prior literature has found that the clustering coefficient and average shortest path in a collaboration network play a promoting role in fostering deep collaboration among team members (Kuperman & Risau-Gusman, 2012) and the efficiency of knowledge dissemination (Wang et al., 2010), which can enhance the overall impact of the team. Therefore, it can be hypothesized that when the ratio of the clustering coefficient to the average shortest path in a research project team's collaboration network increases, the impact of the team's papers will also increase.

Hypothesis 3. (H3): Ratio of clustering coefficient to average shortest path in project team collaboration networks has a positive relationship with impact of outputs.

2.4. The small world of project team collaboration networks and innovation

Small-world networks can act as catalysts for knowledge dissemination and information reciprocity, usher in avant-garde perspectives and methodologies, and serve as a bedrock for innovative research endeavors. On one hand, increasing clustering coefficient can boost innovation outputs because small-world networks allow dense, clustered relationships to coexist with distant and diverse connections. The former can facilitate trust and close collaboration, while the latter can bring entirely novel and non-duplicated information within the clusters (Fleming et al., 2007). On the other hand, decreasing path length also enhances innovation (Cowan & Jonard, 2004) because shorter path lengths enable information to be transmitted to target nodes more quickly, facilitating the exchange of knowledge and ideas. However, overly closed networks perhaps limit information transmission, restricting the input of external viewpoints, and thereby constraining innovative thinking. Chen and Guan (2010) elucidated an inverse U-shaped connection between small-world characteristics and innovation. This insinuates that small-world networks are generally favorable for innovation, but only within a specific range, beyond which the effect reverses.

A noticeable equilibrium exists in a collaborative network, requiring attention to both the clustering coefficient and average shortest path. When the ratio of these two indices skews either excessively high or low, the innovative prowess of a team's outputs tends to wane. A high ratio indicates an occlusive network, inhibiting the introduction of external information and keeping internal knowledge relatively limited. A low ratio suggests a fragmented network, with team members lacking sufficient cooperation and communication, preventing them from deeply exploring and integrating their respective knowledge. Both scenarios act as impediments to innovation. However, when this ratio strikes a middle ground, innovation peaks. At this point, the collaboration network possesses the potential for extensive information dissemination, effective knowledge integration, and cross-group cooperation, contributing to stimulating innovative thinking and resulting in innovative outputs. Given that innovation is gauged via the dual lenses of novelty and conventionality, this investigation posits the subsequent hypothesis:

Hypothesis 4a. **(H4a): There is an inverted U-shaped relationship between the ratio of clustering coefficient to average shortest path length in project team collaboration networks and the novelty of the team's output papers.**

Hypothesis 4b. **(H4b): There is a U-shaped relationship between the ratio of clustering coefficient to average shortest path length in project team collaboration networks and the conventionality of the team's output papers.**

2.5. Interaction between structural holes and small world

Scholars hold three opinions over the interactive influence between structural holes and small world on research performance. Some believe that small world expands the theory of structural holes. In contrast to a purely structural hole perspective, the small-world perspective can assess the optimal configuration of ties within cohesive subgroups and bridging ties (Verspagen & Duysters, 2004). Zhou and Zhao (2011) pointed out that small world can effectively integrate the perspective of social capital and structural holes, providing a better understanding of network efficiency and enterprise innovation performance from a holistic perspective.

Some scholars argue that structural holes complement the limitations of small-world networks. One limitation results from the simplifying assumption that all individuals are connected with no disconnected parts. However, this assumption tends to be violated in many large social networks (Gulati et al., 2012). Chang et al. (2023) found that structural holes can act as a knowledge integrator by bridging diverse good ideas and knowledge from different subgroups. This is an important piece missing from the literature on small-world networks.

Other scholars suppose a complementary effect exists between structural holes and small world. Almahendra (2023) posited an interaction effect between inter-network connections and intra-network cohesion on knowledge acquisition performance in R&D collaborations. This insight into the complementarity between two configurations creates opportunities to enhance the position of R&D companies in the network, which implies that companies can optimize their advantages to be more innovative. Based on the previous discussion, structural holes and small world may exert an interactive influence on the impact and innovation. Thus, the following hypothesis is proposed:

Hypothesis 5. (H5): Structural holes and small world in project team collaboration networks can exert an interactive **influence on the impact and innovation of team's output papers.**

3. Data and research methodology

3.1. Data

The study dataset was obtained from SciSciNet, an open-data repository dedicated to scientific research. It includes over 134 million scientific publications with external references linking to funding sources and public applications (Lin et al., 2023). The National Science Foundation (NSF) is a federal agency in the United States that funds basic research and education. It allocates approximately 25 % of all federal support designated for foundational research at American universities and colleges. Within SciSciNet, there is a compilation of 489,446 NSF awards since 1959. This compilation includes connections between 131,545 grants and 1350,915 scientific publications obtained from NSF.gov. Specifically, connections between NSF awards and the primary papers in

Table 1
Steps of data preprocessing and the remaining NSF projects and their associated papers after each step.

Steps for data preprocessing	Number of NSF projects	Number of associated papers	Percentage of projects filtered out
Raw data	148,148	929,258	0
Step 1: Select connections made through precise DOI matching and those integrated from Crossref as our research data.	143,596	807,149	3.07 %
Step 2: Exclude associated papers with >50 authors.	143,338	801,931	0.17 %
Step 3: Exclude NSF projects with missing grant inception dates or the number of citations of the associated papers and those with fewer than 10 associated papers.	21,618	351,550	82.16 %

SciSciNet were established through precise matches based on paper DOI and fuzzy matches utilizing paper metadata. Additionally, links from Crossref between the NSF and scholarly articles were integrated, yielding 305,314 connection records from NSF to primary papers in SciSciNet.

To ensure the accuracy of the linkage between NSF and papers within SciSciNet, we selected connections made through precise DOI matching and those integrated from Crossref as our research data. Given that papers with >50 authors might skew collaboration networks, we excluded them. Concurrently, papers with missing data, such as the grant's inception date or the number of citations of the linked papers, were omitted. Building on this, to eliminate potential randomness arising from a minimal number of associated papers, we discarded grants linked to fewer than 10 papers. Ultimately, this process yielded a dataset comprising 21,618 NSF grants and their 351,550 associated papers, forming the basis of our research dataset, as shown in Table 1. Fig. 1 shows how projects and associated papers are distributed annually. It can be observed that the majority of the projects are distributed between the years 2000 and 2020, and the number of papers associated with each project exhibits a power-law distribution.

3.2. Variables

3.2.1. Dependent variable

We employ citation counts as a proxy for the academic impact of a research paper. It is noteworthy that while citation counts come with their own set of limitations, they remain the predominant and time-tested metric for gauging academic influence (Yan et al., 2020). To provide a comprehensive representation of the impact derived from a specific funding project, we consider both the mean and median citation counts of papers associated with a particular grant, denoted as *Citation_Count_mean* and *Citation_Count_median*, respectively. The calculation formula for *Citation_Count_mean* is as follows:

$$Citation\_Count\_mean = \frac{1}{n}\sum Citation_t$$

Where n represents the aggregate quantity of scholarly works emanating from the project, and $Citation_t$ denotes citation counts for the t^{th} publication. A higher mean or median citation count associated with a particular project indicates a greater academic influence of the output produced by the project teams.

Innovation often involves reevaluating and reorganizing existing knowledge, methods, ideas, and resources (Cui et al., 2023). References serve as the foundation of a paper's knowledge, and their arrangement can reflect the novelty of the paper. Researchers juxtaposed the actual occurrence frequency of pairs of journals cited in papers with the expected occurrence frequency of randomly generated journal pairs (Uzzi et al., 2013). For each journal pair, they computed a z-score. Specifically, they employ the following formula to compute the z-score value:

$$z = (obs - u)/\sigma$$

In the WoS database, a certain frequency for the journal pair is observed, denoted as *obs*. Through an analysis of ten randomly conducted simulations within the citation network, an average value, u and a standard deviation, σ, are derived. Then, the paper's 10th percentile z-score was employed to represent its novelty, while the median z-score epitomized its conventionality. Lower values for both metrics suggest a higher degree of paper novelty. In this study, the z-scores at the 10th percentile and the median are employed to characterize the novelty and conventionality of a paper. Subsequently, by calculating the mean and median of the 10th percentile z-score and those of the median z-score for papers associated with each project, we gauge the novelty and conventionality of the outputs produced by that project. These are respectively represented by *Atyp_10pct_Z_mean, Atyp_10pct_Z_median, Atyp_Median_Z_mean*, and *Atyp_Median_Z_median*.

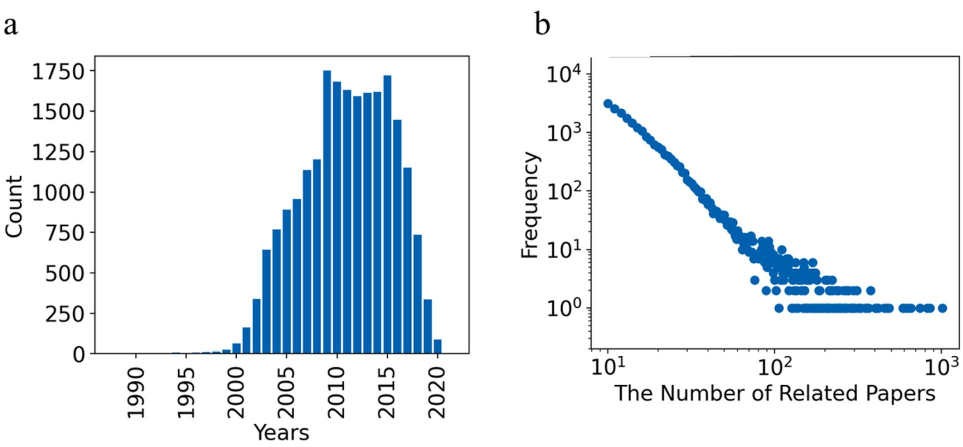

Fig. 1. (a)Annual distribution of NSF projects; (b) Distribution of the number of related papers of NSF projects.

3.2.2. Independent variables

Each project is taken as a unit, and all the papers associated with each project are obtained. We construct corresponding collaboration networks for each project based on the collaboration relationships of team members in the project's papers. Specifically, the network for each project consists of nodes representing the members of the same project team and edges representing collaborations (i.e., co-authorship in the same paper). This network focuses on the interpersonal relationships within the research team, highlighting how team members interact and collaborate. Fig. 2 illustrates the construction process of the collaboration network for project teams. Assume a project results in three papers: paper1, paper2, and paper3. By integrating the collaborative relations from the three papers, the collaboration network corresponding to this project team is formed. Several key network measures are employed based on the collaboration network.

(1) Structural holes

In homogeneous networks, there are differences in information between different clusters, creating a necessity for connections to form structural holes. Nodes that bridge across structural holes connect different groups in the network, forming edges, and they hold intermediary positions that facilitate the spread of information between different groups. In graph theory, a bridge exists if and only if these edges are not part of any cycles. Removing bridges raises the count of graph components. Therefore, bridges have strong structural and information advantages. The number of bridges gauges one's capacity to link diverse social groups, indicating the existence of structural holes in the network (Burt, 2002). We calculate the structural holes of the entire team in every project collaboration network, rather than the sum of each team member's structural holes.

This study employs a bridge-finding algorithm based on chain decompositions to calculate the number of bridges. The core concept involves traversing the entire graph using Depth-First Search (DFS) and simultaneously keeping track of each node's discovery time and minimum discovery time. Here, discovery time represents the timestamp of the first visit to a particular node, while minimum discovery time signifies the lowest discovery time among all nodes reachable from the current node. Bridges are identified by comparing the discovery times and minimum discovery times of the current node and its adjacent nodes. The specific steps are as follows:

(a). For the current node u, examine all adjacent nodes v; (b). If the minimum discovery time of an adjacent node v is greater than the discovery time of the current node u, then the edge between u and v is a bridge. This condition implies that if v can only reach nodes with discovery times earlier than u through u, then the edge between u and v is the only connection between two subgraphs, and its removal would split the graph.

(2) Small world

The study employs a methodology to evaluate the small-world characteristics, following the framework proposed by Fleming et al. (2007). The calculation involves the ratio of the clustering coefficient to the average shortest path length. This ratio effectively summarizes the small-world property by balancing the local node interconnectivity (clustering coefficient) with the network's overall navigability (average shortest path length). This balance is crucial for understanding the efficiency of information flow within collaboration networks.

The clustering coefficient denotes the likelihood of reciprocal connections among the neighboring nodes of a given node (Holland & Leinhardt, 1971). It measures the extent of collaboration among nodes in the collaboration network and is divided into two types: global clustering coefficient and local clustering coefficient. This study utilizes the average clustering coefficient to gauge the overall clustering tendency of the network (Watts &, Strogatz,1998). It represents the mean of the local clustering coefficients across all nodes in the network and can be calculated using the following formula:

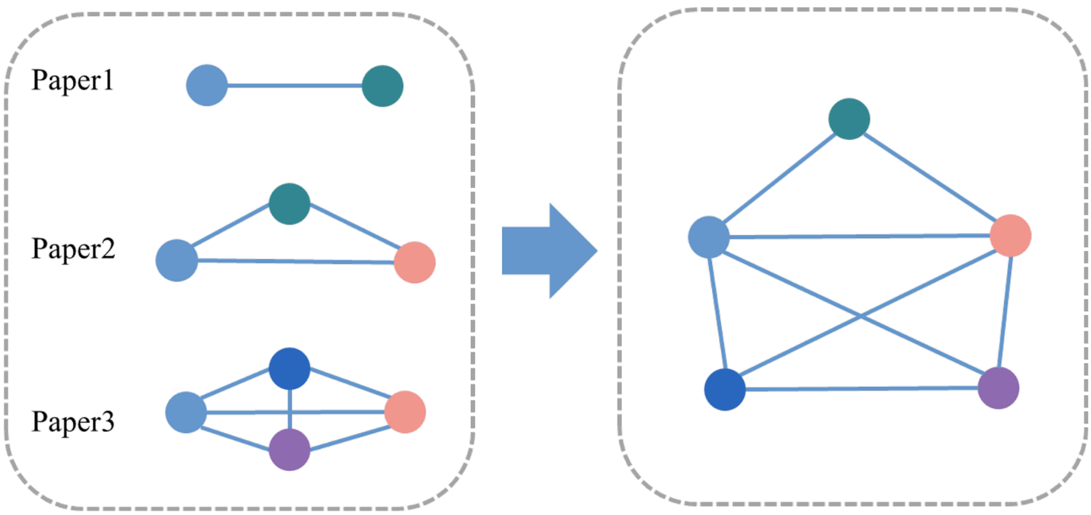

Fig. 2. Illustrative example of the collaboration network construction for project teams. Each node in the figure represents an author of a paper, with the same color indicating the same author. Edges signify collaborative relationships between authors.

$$C_i = \frac{2 \cdot E_i}{k_i \cdot (k_i - 1)}$$

where C_i signifies the local clustering coefficient of node i, E_i represents the count of edges among the neighboring nodes of node i, and k_i denotes the total number of adjacent nodes of node i.

$$\overline{C} = \frac{1}{n} \sum_{i=1}^{n} C_i$$

where $\overline{C}$ stands for the average clustering coefficient of the network, n represents the total node count in the network, and C_i signifies the local clustering coefficient of node i.

The average shortest path length measures the network's efficiency in information propagation by calculating the average distance between nodes. This study employs the Dijkstra algorithm (Dijkstra, 1959) for this calculation.

$$APL = \frac{\sum_{i=1}^{n} \sum_{j \neq i}^{n} d(i,j)}{n \cdot (n-1)}$$

where n is the total node count, and $d(i,j)$ is the shortest path distance between node i and node j.

Therefore, the formula for the ratio of clustering coefficient to average shortest path length (small-worldness coefficient) is as follows:

$$Small\_Worldness = \frac{\overline{C}}{APL}$$

3.2.3. Control variables

To control for confounding factors, the regression model of this study included variables like project inception year, paper count, average author and institution count per paper, average reference count, network node count, and density. Key variables are detailed in Table 2.

3.2.4. Model specification and estimation

In this investigation, we intend to employ a linear regression to validate the aforementioned hypotheses. Specifically, drawing upon the research hypotheses and associated variables, we construct the subsequent regression equation.

$$y = \alpha + \beta Structural\_Holes + \delta Small\_Worldness + \theta Con + \varepsilon$$

Within the specified model, $y$represents the dependent variable, encompassing three distinct dimensions: *AcademicImpact*,*Novelty*, and*Conventionality*. *Structural_Holes*represents the quantity of structural gaps within the project team's collaboration

Table 2
Variable definitions.

Variables (abbr.)	Description
Dependent variables	
Novelty (Atyp_10pct_Z_mean)	Continuous variable; The mean of the 10th percentile z-score for papers associated with each project.
Novelty (Atyp_10pct_Z_median)	Continuous variable; The median of the 10th percentile z-score for papers associated with each project.
Conventionality (Atyp_Median_Z_mean)	Continuous variable; The mean of the median z-score for papers associated with each project.
Conventionality (Atyp_Median_Z_median)	Continuous variable; The median of the median z-score for papers associated with each project.
Academic Impact (Citation_Count_mean)	Continuous variable; The mean of the citation counts for papers associated with each project.
Academic Impact (Citation_Count_median)	Continuous variable; The median of the citation counts for papers associated with each project.
Independent variable	
Structural_Holes	Continuous variable; The number of bridges of collaboration network for project team.
Small_Worldness	Continuous variable; The collaboration network's ratio of clustering coefficient to average path length for a project team.
Control variables	
Nodes	Continuous variable; The number of nodes of collaboration network for project team.
Density	Continuous variable; The density of nodes of collaboration network for project team.
Related_Papers	Continuous variable; The number of papers associated with each project.
Citation_lag_mean	Continuous variable; The mean of the average citation lag for all papers associated with a given project.
Reference_Count_mean	Continuous variable; The mean count of references across all papers associated with a given project.
Team_Size_mean	Continuous variable; The mean count of authors across all papers associated with a given project.
Institution_Count_mean	Continuous variable; The mean count of distinct institutions across all papers associated with a given project.
Years	Continuous variable; The duration in years from each project's start year to 2021.

network.*Small_Worldness* describes the small-world characteristics of the scientific team's collaborative network in the project, suggesting its intricate interlinkages and efficient information transfer potential. The coefficients β and δrespectively delineate the influence of *Structural_Holes* and *Small_Worldness*on the dependent variables, namely *AcademicImpact*,*Novelty*, and *Conventionality*. *Con*encompasses the suite of control variables integrated to account for potential extraneous influences, while$\varepsilon$is indicative of the inherent random error term in the regression construct.

4. Results

4.1. Descriptive statistics

The statistics for all variables are displayed in Table 3. The average number of individuals per team is 61. The average density of the network is 0.224, indicating that the collaborative relationships within the project teams are relatively sparse overall. It is evident that the dependent variables *Atyp_10pct_Z_mean, Atyp_10pct_Z_median, Atyp_Median_Z_mean, Atyp_Median_Z_median, Citation_Count_mean, and Citation_Count_median* all exhibit skewed distributions. To enable ordinary least squares regression analysis, each variable was logarithmically transformed. The average number of structural holes within the project teams is 1.4, and the mean small-worldness coefficient is 0.472.

4.2. The effects of structural holes and small world on scientific impact

Table 4 delineates how *Structural_Holes* and *Small_Worldness* are related to the *Citation_Count_mean*. In Model 1, the coefficient for the variable *Structural_Holes* is -0.0167 ($p < 0.01$), which implies that for every one-unit increase in *Structural_Holes*, the *Citation_Count_mean* decreases by 0.0167 units. This indicates that the academic impact of a project team's published papers negatively correlates with the number of structural holes in their collaboration network.

According to Model 2, the coefficient for the variable *Small_Worldness* is 0.295 ($p < 0.01$), which indicates that a one-unit increase in *Small_Worldness* is associated with a 0.295 unit increase in the *Citation_Count_mean*. This suggests that there exists a positive correlation between small world within the project team and the academic impact of the team's scholarly output. Furthermore, as indicated in Model 3, when both *Structural_Holes* and *Small_Worldness* are included in the regression simultaneously, the results obtained are in line with those obtained from Models 1 and 3. Therefore, Hypotheses 1 and 3 are corroborated.

To further assess the presence of multicollinearity among the predictors, variance inflation factors (VIFs) were computed. The findings are delineated in Table S1 in the Appendix. The VIF values for all predictor variables were below 10, suggesting an absence of significant multicollinearity concerns (Huang et al., 2022b). These findings permit the advancement of regression analysis.

4.3. The effects of structural holes and small world on innovation

Table 5 reveals the correlation between *Structural_Holes* and *Small_Worldness* in relation to *Atyp_10pct_Z_mean*. In Model 4, the coefficient value for *Structural_Holes* stands at 0.00736 ($p < 0.01$), suggesting that for each one-unit increase in *Structural_Holes*, the *Atyp_10pct_Z_mean* increases by 0.00736 units. This denotes a direct linkage between the number of structural holes within a project team and the novelty of the team's academic publications. Consequently, H2a does not hold true.

Model 5 shows a coefficient value of -0.454 for *Small_Worldness* and a coefficient of 1.958 for the square of *Small_Worldness*, both significant at the 0.01 level. These values denote a U-curve correlation between *Small_Worldness* and *Atyp_10pct_Z_mean*, implying an inverse U-curve correlation between small world within the project team and novelty of the team's research output. This suggests that

Table 3
Summary of statistical data.

Variables	Obs	Mean	Std. Dev.	Min	Max
Atyp_10pct_Z_mean	21,618	31.371	133.659	−55.712	3787.488
Atyp_10pct_Z_median	21,618	2.29	34.012	−49.597	1078.642
Atyp_Median_Z_mean	21,618	242.728	412.161	−0.143	5416.122
Atyp_Median_Z_median	21,618	170.577	293.049	−0.964	4564.701
Citation_Count_mean	21,618	49.259	66.499	0.417	3308.462
Citation_Count_median	21,618	28.689	30.458	0	606
Structural_Holes	21,618	1.4	3.69	0	247
Small_Worldness	21,618	0.472	0.068	0.049	1
Nodes	21,618	61.001	102.504	4	3716
Density	21,618	0.224	0.112	0.002	1
Related_Papers	21,618	21.023	28.84	10	1014
Reference_Count_mean	21,618	49.841	19.473	6.231	198.706
Team_Size_mean	21,618	5.175	2.502	1.167	43.667
Institution_Count_mean	21,618	2.51	1.322	1	17
Years	21,618	10.046	4.468	0	33
Citation_lag_mean	21,618	3.577	1.877	−0.48	11.37

Note: Due to the inclusion of a small number of preprint citations, the minimum value of Citation_lag_mean is <0.

Table 4
OLS on *Citation_Count_mean*, with robust standard errors in parentheses.

Variable	Citation_Count_mean		
	Model 1	Model 2	Model 3
Structural_Holes	−0.0167***		−0.0161***
	(0.00201)		(0.00194)
Small_Worldness		0.295***	0.219***
		(0.0694)	(0.0694)
Nodes	−0.000676***	−0.000380***	−0.000668***
	(0.000119)	(0.000124)	(0.000117)
Density	−0.274***	−0.336***	−0.322***
	(0.0423)	(0.0453)	(0.0448)
Related_Papers	0.00467***	0.00226***	0.00456***
	(0.000462)	(0.000447)	(0.000458)
Reference_Count_mean	0.0120***	0.0121***	0.0119***
	(0.000272)	(0.000273)	(0.000272)
Team_Size_mean	0.0414***	0.0434***	0.0400***
	(0.00361)	(0.00363)	(0.00362)
Institution_Count_mean	0.000508	−0.00402	0.00161
	(0.00557)	(0.00557)	(0.00557)
Years	−0.00234	−0.00224	−0.00224
	(0.00248)	(0.00249)	(0.00248)
Citation_lag_mean	0.351***	0.350***	0.351***
	(0.00603)	(0.00604)	(0.00603)
_cons	1.482***	1.365***	1.397***
	(0.0247)	(0.0357)	(0.0359)
N	21,618	21,618	21,618
R^2	0.492	0.490	0.492
adj. R^2	0.491	0.489	0.492

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5
OLS on *Atyp_10pct_Z_mean*, with robust standard errors in parentheses.

Variable	Atyp_10pct_Z_mean		
	Model 4	Model 5	Model 6
Structural_Holes	0.00736***		0.00515**
	(0.00239)		(0.00209)
Small_Worldness		−0.454***	−0.435***
		(0.0715)	(0.0716)
Small_Worldness squared		1.958***	1.863***
		(0.417)	(0.420)
Nodes	−0.0000664	−0.000213*	−0.000120
	(0.000133)	(0.000121)	(0.000128)
Density	0.0728*	0.129***	0.128***
	(0.0405)	(0.0434)	(0.0436)
Related_Papers	0.0000443	0.00115***	0.000416
	(0.000437)	(0.000390)	(0.000434)
Reference_Count_mean	−0.00339***	−0.00324***	−0.00321***
	(0.000264)	(0.000264)	(0.000264)
Team_Size_mean	−0.0436***	−0.0410***	−0.0399***
	(0.00388)	(0.00365)	(0.00378)
Institution_Count_mean	0.184***	0.182***	0.181***
	(0.00591)	(0.00581)	(0.00588)
Years	0.0326***	0.0321***	0.0321***
	(0.00236)	(0.00236)	(0.00236)
Citation_lag_mean	−0.0574***	−0.0565***	−0.0567***
	(0.00535)	(0.00534)	(0.00534)
_cons	3.974***	3.933***	3.934***
	(0.0260)	(0.0258)	(0.0262)
N	21,618	21,618	21,618
R2	0.090	0.094	0.094
adj. R2	0.089	0.093	0.094

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

either exceedingly low or high levels of small world might stifle the team's innovative prowess. Additionally, Model 6 incorporates both *Structural_Holes* and *Small_Worldness* along with *Small_Worldness squared* into the regression. The outputs corroborate those of Models 4 and 5. Therefore, H4a is confirmed.

Table 6 shows how *Structural_Holes, Small_Worldness*, and *Atyp_Median_Z_mean* are interconnected. In Model 7, the value of the coefficient associated with *Structural_Holes* stands at −0.00245. Although there is no notable relationship between the count of structural holes in a project team and the conventionality of their scholarly publications, the results do show a shift in Model 9. When *Structural_Holes, Small_Worldness*, and *Small_Worldness squared* are considered in the regression analysis, H2b is confirmed. This variation underscores the importance of model specifications and the multifaceted nature of the relationships at hand.

Turning to Model 8, the coefficient for *Small_Worldness* is −0.504 ($p < 0.01$), and *Small_Worldness squared* has a coefficient of 1.536 ($p < 0.01$). These coefficients point to a U-curve correlation between small world and the conventionality of the team's academic output, signifying that moderate levels of small world can facilitate the team's ability to produce innovative research. Moreover, as demonstrated in Model 9, when *Structural_Holes, Small_Worldness*, and *Small_Worldness squared* are all included in the regression analysis, the results align well with those from Model 8. Therefore, H4b is substantiated.

To validate the non-linear link between small world in project teams and the novelty of their outputs, we illustrated their curves graphically. As illustrated in Fig. 3, the data further corroborates an inverted U-shaped relationship between small world within team collaboration networks and the novelty of the research publications produced by the team.

4.4. The interaction between structural holes and small world

Table 7 displays the moderating effect of structural holes on the relationship between small world and research outcomes. Specifically, column (1) reveals that the interaction term between *Structural_Holes* and *Small_Worldness* is insignificant, indicating that structural holes do not moderate the relationship between small world and impact. Columns (2) and (3) show that both *Structural_Holes* and *Small_Worldness squared* are insignificant, suggesting that structural holes also do not moderate the relationship between small world and innovation. Consequently, Hypothesis 5 is not supported.

We also examined the moderated effect of small world on the relationship between structural holes and research outcomes, as shown in Table S2 in the appendix. Specifically, the coefficients of the interaction terms between *Small_Worldness* and *Structural_Holes* in columns (1) to (3) are not significant, implying that small world does not moderate the relationship between structural holes and research outcomes. This also corroborates that Hypothesis 5 is not supported.

4.5. The spillover effects of interdisciplinary project collaborations

Furthermore, we examine the spillover effects of interdisciplinary project collaborations by analyzing the citations of project papers by those from different disciplines. Specifically, the MAG database categorizes research papers into multiple fields, forming a six-level hierarchical structure (i.e., Levels 0, 1, 2, 3, 4, and 5). At the highest tier, Level 0 encompasses 19 macro-level disciplines frequently

Table 6
OLS on *Atyp_Median_Z_mean*, with robust standard errors in parentheses.

Variable	Atyp_Median_Z_mean		
	Model 7	Model 8	Model 9
Structural_Holes	−0.00245		−0.00478**
	(0.00245)		(0.00223)
Small_Worldness		−0.504***	−0.521***
		(0.102)	(0.102)
Small_Worldness squared		1.536***	1.623***
		(0.534)	(0.535)
Nodes	−0.000571***	−0.000538***	−0.000624***
	(0.000176)	(0.000186)	(0.000174)
Density	0.298***	0.376***	0.377***
	(0.0616)	(0.0664)	(0.0663)
Related_Papers	0.00311***	0.00283***	0.00351***
	(0.000637)	(0.000649)	(0.000642)
Reference_Count_mean	0.0000998	0.000326	0.000293
	(0.000389)	(0.000390)	(0.000390)
Team_Size_mean	−0.00481	0.000232	−0.000773
	(0.00426)	(0.00420)	(0.00421)
Institution_Count_mean	0.191***	0.186***	0.188***
	(0.00818)	(0.00811)	(0.00817)
Years	0.0291***	0.0286***	0.0286***
	(0.00341)	(0.00341)	(0.00341)
Citation_lag_mean	−0.0693***	−0.0689***	−0.0686***
	(0.00785)	(0.00783)	(0.00783)
_cons	4.397***	4.351***	4.350***
	(0.0356)	(0.0368)	(0.0363)
N	21,618	21,618	21,618
R2	0.070	0.072	0.072
adj. R2	0.070	0.072	0.072

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

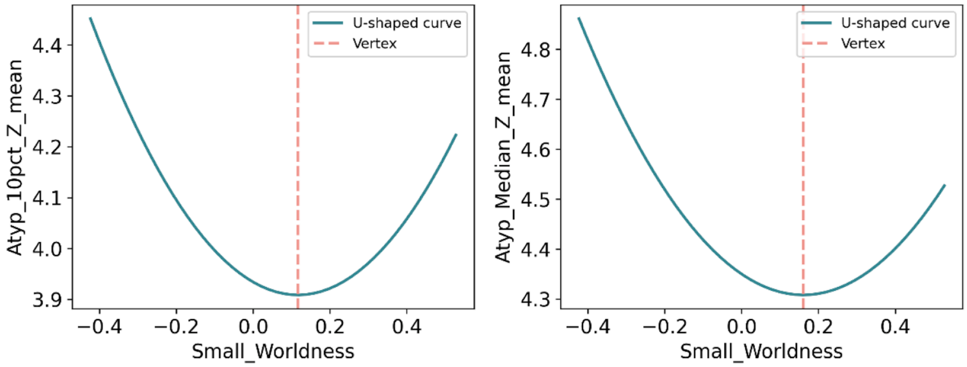

Fig. 3. The U-curve influence of *Small_Worldness* on *Atyp_10pct_Z_mean* and *Atyp_Median_Z_mean.*

Table 7
The moderated effect of structural holes between small world and research outcomes, with robust standard errors in parentheses.

Variable	(1) Citation_Count _mean	(2) Atyp_10pct _Z_mean	(3) Atyp_Median _Z_mean
Small_Worldness	0.244***	−0.515***	−0.603***
	(0.0803)	(0.0806)	(0.114)
Structural_Holes	−0.0164***	0.00609***	−0.00388
	(0.00191)	(0.00213)	(0.00239)
Small_Worldness squared		1.984***	1.704***
		(0.516)	(0.658)
Structural_Holes*Small_Worldness	−0.0104	0.0353*	0.0368*
	(0.0150)	(0.0183)	(0.0212)
Structural_Holes*Small_Worldness squared		−0.000578	0.0134
		(0.0976)	(0.120)
Nodes	−0.000664***	−0.000136	−0.000641***
	(0.000117)	(0.000130)	(0.000176)
Density	−0.327***	0.142***	0.392***
	(0.0454)	(0.0442)	(0.0674)
Related_Papers	0.00455***	0.000427	0.00352***
	(0.000458)	(0.000436)	(0.000643)
Reference_Count_mean	0.0119***	−0.00321***	0.000293
	(0.000272)	(0.000264)	(0.000390)
Team_Size_mean	0.0399***	−0.0395***	−0.000329
	(0.00362)	(0.00379)	(0.00423)
Institution_Count_mean	0.00167	0.180***	0.187***
	(0.00557)	(0.00588)	(0.00817)
Years	−0.00221	0.0321***	0.0285***
	(0.00248)	(0.00236)	(0.00341)
Citation_lag_mean	0.351***	−0.0567***	−0.0686***
	(0.00603)	(0.00534)	(0.00784)
_cons	1.501***	3.930***	4.346***
	(0.0254)	(0.0262)	(0.0364)
N	21,618	21,618	21,618
R2	0.492	0.094	0.073
adj. R2	0.491	0.094	0.072

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

used in disciplinary classification (Huang et al., 2022a). Each paper is assigned to a corresponding discipline, allowing us to determine whether the citing and cited papers belong to the same discipline. The citations a paper receives from other disciplines are then used as a new dependent variable for regression analysis. The results are presented in Table 8.

Models (1) and (2) use *Citation_Count_mean* and *Citation_Count_median* as dependent variables, respectively. The results indicate that *Structural_Holes* is negatively correlated with both *Citation_Count_mean* and *Citation_Count_median* ($p < 0.01$), with coefficients of −0.00979 and −0.00780, respectively. This demonstrates that an increase in *Structural_Holes* by one unit is associated with a decrease of 0.00979 units in *Citation_Count_mean* and a decrease of 0.00780 units in *Citation_Count_median. Small_Worldness* is positively correlated with both *Citation_Count_mean* and *Citation_Count_median* ($p < 0.01$), with coefficients of 0.825 and 0.609, respectively. This implies that an increase in *Small_Worldness* by one unit is associated with an increase of 0.825 units in *Citation_Count_mean* and 0.609 units in *Citation_Count_median*. Model (3) also uses *Citation_Count_median* as the dependent variable and employs negative binomial regression for analysis, which yields consistent results. These results suggest that the *Structural_Holes* and *Small_Worldness* of

Table 8
OLS and NBR on *Citation_Count_mean, Citation_Count_ median* (citations by other disciplines), with robust standard errors in parentheses.

Variable	(1) Citation_Count_ mean	(2) Citation_Count_ median	(3) Citation_Count_ median
Structural_Holes	−0.00979***	−0.00780***	−0.00849***
	(0.00235)	(0.00214)	(0.00276)
Small_Worldness	0.825***	0.609***	0.602***
	(0.0980)	(0.0830)	(0.109)
Nodes	0.000218	0.0000882	0.0000355
	(0.000186)	(0.000140)	(0.000175)
Density	−0.869***	−0.303***	−0.0429
	(0.0633)	(0.0534)	(0.0692)
Related_Papers	0.00163***	0.000727	0.000347
	(0.000613)	(0.000494)	(0.000621)
Reference_Count_mean	0.00428***	0.00373***	0.00554***
	(0.000383)	(0.000322)	(0.000440)
Team_Size_mean	0.0416***	0.0304***	0.0325***
	(0.00505)	(0.00402)	(0.00558)
Institution_Count_mean	−0.0927***	−0.0812***	−0.0494***
	(0.00816)	(0.00671)	(0.00921)
Years	−0.0409***	−0.0419***	−0.0529***
	(0.00352)	(0.00311)	(0.00417)
Citation_lag_mean	0.326***	0.292***	0.392***
	(0.00851)	(0.00759)	(0.0100)
_cons	1.322***	0.987***	0.686***
	(0.0508)	(0.0432)	(0.0569)
N	21,618	21,618	21,618
R^2	0.213	0.203	
adj. R^2	0.212	0.203	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

interdisciplinary collaboration networks have spillover effects on interdisciplinary knowledge citation and flow. In particular, the small world of interdisciplinary collaboration networks can facilitate the spillover effect of interdisciplinary knowledge, while structural holes tend to inhibit this effect.

4.6. Robustness checking

To further validate the findings of this study, we conducted five different robustness checks. First, we substituted the dependent variables for impact, novelty, and conventionality from mean values (*Citation_Count_mean, Atyp_10pct_Z_mean, Atyp_Median_Z_mean*) to median values (*Citation_Count_median, Atyp_10pct_Z_median, Atyp_Median_Z_median*). The outputs are presented in Table S3 in the appendix. The findings largely align with our previous conclusions: A negative relationship exists between the number of structural holes in a project team's collaboration network and both the impact and conventionality of the papers produced, suggesting that while structural holes can enhance the novelty of the research output, they may diminish its impact. Conversely, small world in the project team's collaboration network positively correlates with the research papers' impact and displays an inverted U-curve relationship with their novelty.

Second, the clustering coefficient and average path length are normalized by a random network to remeasure *Small_Worldness*, replacing the previous independent variable (Uzzi & Spiro, 2005). A new regression analysis was then conducted. Table S4 in the appendix shows that the results remained consistent with the previous findings, indicating good robustness.

Third, for the count-type dependent variable, *Citation_Count_median*, whose variance significantly exceeds its mean, a negative binomial regression was employed to reanalyze for robustness testing. As shown in Table S5, the results remained consistent, indicating that structural holes are negatively correlated with academic impact, while small world is positively correlated with academic impact.

Furthermore, to mitigate the impact of self-citations on the results, we incorporated the pattern of self-citations into the study. This was determined based on whether there was an overlap between the authors of the citing and cited papers. After excluding self-citations, we re-analyzed the relationship between *Structural_Holes, Small_Worldness*, and *Citation_Count_mean* and *Citation_Count_median*. The results, as presented in Table S6 in the appendix, remained consistent, indicating that the findings of this study exhibit good robustness.

Finally, we investigate the dynamics of evolving networks by comparing measurements from a series of static snapshots over time. Among the projects analyzed, 99 % of the papers associated with them have an average interval of <15 years between their initiation and publication. Consequently, we selected 15 years as the time span for our evolutionary analysis, further dividing it into three periods: 0–5 years, 0–10 years, and 0–15 years. Networks were constructed for each period, and regression analyses were conducted. The results, as shown in the appendix, generally align with those obtained in earlier phases of the study, demonstrating consistency as the networks evolved.

5. Discussion and conclusion

This study constructs a scientific collaboration network from the perspective of project teams, figuring out the effect of two network features, structural holes and small world, on the impact and innovation of scientific outputs. Our findings indicate a negative correlation between structural holes and impact, but a positive correlation with innovation. Small world positively correlates with impact but exhibits an inverted U-shaped relationship with innovation. Furthermore, there is no significant interaction between structural holes and small world in their effects on impact and innovation within the network. It is also found that there are interdisciplinary spillover effects of project collaborations.

5.1. Theoretical implications

Previous research has mainly focused on the influence of collaboration networks based on individual research papers on scientific performance. In contrast, we construct a scientific collaboration network based on project teams, offering a fresh source for understanding scientific collaboration network structures. Meanwhile, we include two network structural features closely related to information flow, structural holes, and small world, into our analytical framework. This complements related research on the relationship between team knowledge dissemination and scientific impact and innovation.

In project collaboration networks, the number of structural holes inhibits the impact of the team's outputs, which aligns with the perspective of Li et al. (2017). This finding suggests that if there are too many members spanning structural holes in project teams, it may lead to more social friction, reduce the efficiency of information transmission, and adversely affect the overall impact. While small world promotes the impact, compensating for the shortcomings of structural holes. Therefore, establishing close connections among team members can mitigate cognitive conflicts.

Structural holes facilitate innovation of project team's outputs, which is partially consistent with the findings of Wang et al. (2023). However, the difference lies in that they understood innovation from the perspective of novelty. They found that scientists bridging structural holes produce more novel research. In contrast, we measure innovation from both the dimensions of novelty and conventionality, finding that structural holes weaken the conventionality of outputs. This suggests that in project collaboration networks, the advantage of heterogeneous knowledge brought by structural holes can break traditional thinking patterns, promote the integration of knowledge from diverse subgroups, and stimulate innovative viewpoints and outputs. Especially by forming project teams with members from different backgrounds, it is possible to create key connection nodes within the network, allowing the integration of expertise and perspectives from various subgroups. This diversity can avoid the limitations of single-track thinking, forming bridges for information and resources, and fostering the development of innovative solutions. While the relationship between small world and innovation is more complex, showing an inverted U-shaped relationship. This finding complements the general cognition of a positive relationship in existing research (Galaso & Kovarik, 2021). It indicates that the dense cluster ensures rapid sharing of internal knowledge and efficient development of new ideas, but overly closed networks hinder the influx of new knowledge. Therefore, in a small-world network, it is advisable to have an appropriate number of members bridging structural holes to enrich the knowledge heterogeneity within the team, ensuring the innovation foundation.

We found that project-based collaboration networks exhibit disciplinary spillover effects, supporting the viewpoint of Zou et al. (2023). This indicates that such networks are more effective in promoting cross-disciplinary knowledge flow and expanding the impact of project outcomes. This effect further explains the role of structural holes and small-world phenomena within academic collaboration networks, providing new insights into the deeper functions of network characteristics in research collaboration environments and contributing to the development of network theory. The increase in cross-disciplinary citations means that an appropriate number of structural holes and a tightly-knit small-world network can quickly adapt to and integrate knowledge and methods from different disciplines, thereby enhancing the impact of project outcomes across multiple fields.

5.2. Practical implications

Our findings offer valuable insights for project team management and collaboration. Firstly, when assembling a project team, it is necessary to recruit a moderate number of team members who cross disciplinary and organizational boundaries. These scholars spanning structural holes can introduce rich heterogeneous knowledge and social capital to the team, enhancing the potential for effective knowledge recombination and aiding the project team in achieving greater innovation.

Secondly, project teams should devise collaborative strategies based on their innovation objectives. The inverted U-shaped relationship between small world and innovation provides different insights for various types of projects. For scientific projects committed to exploratory innovation, team members should establish deep collaborations to ensure efficient knowledge dissemination within the team while also ensuring that the small-world network is not overly closed. This is because exploratory innovation requires high-quality collaborative relationships to absorb diverse knowledge (Ma et al., 2020). And structural holes can promote the acquisition of heterogeneous information by connecting unrelated collaborators. Therefore, within the team, members spanning structural holes can organize open discussions to reduce cognitive conflicts among subgroups with different backgrounds in the team and facilitate the effective integration of heterogeneous knowledge. Outside the project team, the leaders of projects should leverage their role advantages by occupying structural holes within the wider collaboration network and actively seeking brokerage opportunities. The specific way involves participating in broader research communities to accumulate more academic capital, thereby shaping comprehensive understanding conducive to their projects.

For enterprise projects committed to exploitative innovation, companies have a high degree of knowledge anchoring, which means an emphasis on reutilizing existing organizational knowledge (Benner & Tushman, 2003). This indicates that companies focus on the efficiency of developing new products, requiring high-quality knowledge within project teams and close collaboration among team members. Additionally, small world can contribute to knowledge protection within project teams. There are risks of exploitation when collaborating with central leading companies in profit-oriented market competition (Ma et al., 2020), and collaborating with partners possessing heterogeneous knowledge may lead to conflicts and coordination issues (Jarvenpaa & Välikangas, 2014), thus constraining corporate innovation. Therefore, in such project collaboration networks, companies should focus on optimizing collaboration among internal team members. When forming collaborative relationships between companies, it is suggested to carefully select partners, such as upstream or downstream companies in the industry chain that do not pose competitive threats but have a solid foundation of knowledge in the relevant domain.

In addition, for intra-cluster links, regular team meetings and workshops can be established, project knowledge bases can be built, and specialized project management software can be used to promote knowledge sharing and idea exchange among members, ensuring efficient communication and collaboration. For inter-cluster links, new knowledge and resources can be acquired, and influence and innovativeness can be expanded through actively engaging in external academic exchanges, participating in industry alliances, and joining open innovation platforms.

Thirdly, the finding that structural holes and small world operate independently in influencing impact and innovation also has practical and policy implications. Funding agencies and institutions can design policies to foster the formation of both structural holes and small-world networks within research collaborations, such as establishing interdisciplinary research awards, supporting cross-institutional partnerships, implementing generational mentoring, and creating collaborative team spaces. The dual focus can help in achieving a balance between innovation and impact, ultimately leading to more effective and productive research environments.

Lastly, the cross-disciplinary spillover effect of project collaboration is reflected in broader academic citations and dissemination, which can help enhance the interdisciplinary impact of project outcomes. Funding agencies and institutions can develop more targeted funding strategies, such as prioritizing interdisciplinary projects with high spillover effects to optimize resource allocation and improve the effectiveness of research investments.

5.3. Limitations and future work

This study has its limitations. Firstly, the conclusions drawn are primarily applicable to the scenario using the NSF as the research sample. The generalizability to other project teams remains to be further verified. Subsequent research might consider selecting teams from other funding institutions for a comparative analysis. Secondly, despite incorporating numerous control variables, the study could still be influenced by unobserved factors, such as the amount of project funding or the academic standing of the project leader. Due to data accessibility constraints, not all potential variables could be controlled for. Future investigations should contemplate introducing additional control variables and additional factors. Moreover, this research does not dissect the variances across different academic disciplines, and ensuing studies might delve deeper into the heterogeneity present across these fields. Despite certain limitations, this study provides insightful implications for the effective construction of project teams. In further research, we will explore how to optimize the collaboration network structure of project teams to enhance research impact while preserving innovation. This could involve introducing a balanced metric, which measures both impact and innovation. Further studies might also investigate the role of technology in facilitating or hindering the benefits of network features such as structural holes and small world, considering the increasing digitization of research environments.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to improve language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

CRediT authorship contribution statement

Zhifeng Liu: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Chenlin Wang:** Writing – review & editing, Writing – original draft, Conceptualization. **Jinqing Yang:** Writing – original draft, Funding acquisition.

Declaration of competing interest

The authors report no declarations of interest.

Acknowledgments

This work was supported by Youth Program of the National Natural Science Foundation of China (No. 72304108) and Basic Research Program of Chinese Universities (No. CCNU23XJ013).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.joi.2024.101611.

References

Almahendra, R. (2023). Disentangling learning network dilemma: A small-world effect in a globalized world. *Sustainability, 15*(3), 2288.

Andrikopoulos, A., Bekiaris, M., & Kostaris, K. (2020). Stars in a small world: Social networks in auditing research. *Scientometrics, 122*(1), 625–643.

Benner, M. J., & Tushman, M. L. (2003). *Exploitation, exploration, and process management: The productivity dilemma revisited, 28* pp. 238–256). Academy of Management.

Benz, P., & Rossier, T. (2022). Is interdisciplinarity distinctive? Scientific collaborations through research projects in natural sciences. *Social Science Information, 61*(1), 179–214.

Brunetta, F., Boccardelli, P., & Lipparini, A. (2015). Central positions and performance in the scientific community. Evidences from clinical research projects. *Journal of Business Research, 68*(5), 1074–1081.

Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge: Harvard University Press.

Burt, R. S. (2002). Bridge decay. *Social Networks, 24*(4), 333–363.

Chang, S., Lee, J., & Song, J. (2023). Giant cluster formation and integrating role of bridges in social diffusion. *Strategic Management Journal, 44*(12), 2950–2985.

Chen, Z. F., & Guan, J. C. (2010). The impact of small world on innovation: An empirical study of 16 countries. *Journal of Informetrics, 4*(1), 97–106.

Cowan, R., & Jonard, N. (2004). Network structure and the diffusion of knowledge. *Journal of Economic Dynamics & Control, 28*(8), 1557–1575.

Cui, H., Li, T., & Wang, C. J. (2023). Climbing up the ladder of abstraction: How to span the boundaries of knowledge space in the online knowledge market? *Humanities and Social Sciences Communications, 10*(1), 1–12.

Dai, R., Donohue, L., Drechsler, Q., & Jiang, W. (2023). Dissemination, publication, and impact of finance research: When novelty meets conventionality. *Review of Finance,, 27*(1), 79–141.

De Stefano, D., Fuccella, V., Prosperina, V. M., & Zaccarin, S. (2013). The use of different data sources in the analysis of co-authorship networks and scientific performance. *Social Networks, 35*(3), 370–381.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik, 1*(1), 269–271.

Ebadi, A., & Schiffauerova, A. (2015). On the relation between the small world structure and scientific activities. *PloS one, 10*(3).

Fleming, L., King, C., & Juda, A. (2007). Small worlds and regional innovation. *Organization Science, 18*(6), 938–954.

Fronczak, A., Mrowinski, M. J., & Fronczak, P. (2022). Scientific success from the perspective of the strength of weak ties. *Scientific Reports, 12*(1).

Galaso, P., & Kovarik, J. (2021). Collaboration networks, geography and innovation: Local and national embeddedness. *Papers in Regional Science, 100*(2), 349–377.

Galunic, D. C., & Rodan, S. (1998). Resource recombinations in the firm: Knowledge structures and the potential for schumpeterian innovation. *Strategic Management Journal, 19*(12), 1193–1201.

Guan, J. C., Yan, Y., & Zhang, J. J. (2017a). The impact of collaboration and knowledge networks on citations. *Journal of Informetrics, 11*(2), 407–422.

Guan, J. C., Zhang, J. J., & Yan, Y. (2017b). A dynamic perspective on diversities and network change: Partner entry, exit and persistence. *International Journal of Technology Management, 74*(1–4), 221–242.

Gulati, R., Sytch, M., & Tatarynowicz, A. (2012). The rise and fall of small worlds: Exploring the dynamics of social structure. *Organization Science, 23*(2), 449–471.

Halevy, N., Halali, E., & Zlatev, J. J. (2019). Brokerage and brokering: An integrative review and organizing framework for third party influence. In *Academy of Management Annals, 13* pp. 215–239).

Holland, P. W., & Leinhardt, S. (1971). Transitivity in structural models of small groups. *Comparative Group Studies, 2*(2), 107–124.

Huang, Y., Lu, W., Liu, J., Cheng, Q., & Bu, Y. (2022a). Towards transdisciplinary impact of scientific publications: A longitudinal, comprehensive, and large-scale analysis on Microsoft Academic Graph. *Information Processing & Management, 59*(2), Article 102859.

Huang, Z., Zong, Q., & Ji, X. (2022b). The associations between scientific collaborations of LIS research and its policy impact. *Scientometrics, 127*(11), 6453–6470.

Hung, C. L. (2017). Social networks, technology ties, and gatekeeper functionality: Implications for the performance management of R&D projects. *Research Policy, 46*(1), 305–315.

Jarvenpaa, S. L., & Välikangas, L. (2014). Opportunity creation in innovation networks. *California Management Review, 57*(1), 67–87.

Kim, H., & Park, Y. (2009). Structural effects of R&D collaboration network on knowledge diffusion performance. *Expert Systems with Applications, 36*(5), 8986–8992.

Kuperman, M. N., & Risau-Gusman, S. (2012). Relationship between clustering coefficient and the success of cooperation in networks. *Physical Review E, 86*(1).

Li, M. Z., Zhuang, X. L., Liu, W. X., et al. (2017). More stable ties or better structure? An examination of the impact of co-author network on team knowledge creation. *Frontiers in Psychology, 8*.

Lin, Z., Yin, Y., Liu, L., & Wang, D. (2023). SciSciNet: A large-scale open data lake for the science of science research. *Scientific Data, 10*(1), 315.

Ma, D., Zhang, Y. R., & Zhang, F. (2020). The influence of network positions on exploratory innovation: An empirical evidence from china's patent analysis. *Science Technology and Society, 25*(1), 184–207.

Moody, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963 To 1999. *American Sociological Review, 69*, 213–238.

Mukherjee, S., Uzzi, B., Jones, B., & Stringer, M. (2016). A new method for identifying recombinations of existing knowledge associated with high-impact innovation. *Journal of Product Innovation Management, 33*(2), 224–236.

O'Connor, C. (2021). Measuring conventionality. *Australasian Journal of Philosophy, 99*(3), 579–596.

Park, J., Park, J., & Choi, Y. (2022). The driving force for collaboration networks in environmental engineering in South Korea. *Environmental Engineering Research, 27*(2).

Rose, M. E. (2022). Small world: Narrow, wide, and long replication of Goyal, van der Leij and Moraga-Gonzelez (JPE 2006) and a comparison of EconLit and Scopus. *Journal of Applied Econometrics, 37*(4), 820–828.

Rost, K. (2011). The strength of strong ties in the creation of innovation. *Research Policy, 40*(4), 588–604.

Schillebeeckx, S. J. D., Lin, Y., & George, G. (2019). When do expert teams fail to create impactful inventions? *Journal of Management Studies, 56*(6), 1073–1104.

Su, Y., & Jiang, X. S. (2023). Simulation research on knowledge flow in a collaborative innovation network. *Expert Systems, 40*(7).

Tahmooresnejad, L., Beaudry, C., & Mirnezami, S. R. (2021). The study of network effects on research impact in Africa. *Science and Public Policy, 48*(4), 462–473.

Uzzi, B., Mukherjee, S., Stringer, M., et al. (2013). Atypical combinations and scientific impact. *Science, 342*(6157), 468–472 (New York, N.Y.).

Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The small world problem. *American Journal of Sociology, 111*(2), 447–504.

Verspagen, B., & Duysters, G. (2004). The small worlds of strategic technology alliances. *Technovation, 24*(7), 563–571.

Vieira, E. S. (2023). The influence of research collaboration on citation impact: The countries in the European Innovation Scoreboard. *Scientometrics, 128*, 3555–3579.

Walter, C. E., Ferasso, M., Veloso, C. M., & Auyong-Oliveira, M. (2022). The creation and dissemination of scientific knowledge: An empirical analysis of economic Nobel prizes and open innovation scientific performance. *Quality-Access to Success, 23*(187), 232–243.

Wang, W. P., Shen, Q. Y., & Chen, Y. Q. (2010). Community degree, clustering coefficient and knowledge propagation efficiency in complex networks. *Advances in Grey Systems Research* (pp. 561–569).

Wang, Y., Li, N., Zhang, B., et al. (2023). The effect of structural holes on producing novel and disruptive research in physics. *Scientometrics, 128*(3), 1801–1823.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*(6684), 440–442.

Xiang, Q., Lin, D. M., & Wang, J. J. (2023). The formation, consolidation, and transition of international brokerage networks: The case of an international new venture in an emerging market. *Journal of International Management, 29*(1), Article 100995.

Yan, Y., Tian, S., & Zhang, J. (2020). The impact of a paper's new combinations and new components on its citation. *Scientometrics, 122*, 895–913.

Yang, K. (2022, April 7). Urgent need for in-depth reform of research project organization management [News]. Guangming Daily, (16).

Yang, Y., She, Y., Hong, J., et al. (2021). The centrality and innovation performance of the quantum high-level innovation team: The moderating effect of structural holes. *Technology Analysis & Strategic Management, 33*(11), 1332–1346.

Zaheer, A., & Soda, G. (2009). Network evolution: The origins of structural holes. *Administrative Science Quarterly, 54*(1), 1–31.

Zhou, W., & Li, H. (2023). A study on the multidimensional driving mechanism of cross-regional scientific collaboration network in China. *Technology Analysis & Strategic Management*.

Zhou, W., & Zhao, Y. (2011). Research on small world effect in complex innovation network: From the integrating perspective of social capital and structural hole. In *Proceedings of the 7th China science and technology policy and management academic annual conference* (p. 10).

Zou, B. T., Wang, Y. F., Kwoh, C. K., & Cen, Y. H. (2023). Directed collaboration patterns in funded teams: A perspective of knowledge flow. *Information Processing & Management, 60*(2), Article 103237.