

Can One Uphold Tradition and Innovate?—An Analysis of Novelty and Conventionality in Chinese Library and Information Science Papers (Postprint)

Authors: Liang Xingkun

Date: 2022-11-27T00:00:00+00:00

Abstract

[Purpose/Significance] Taking China' s library and information science (LIS) field as an example, this study measures the novelty and conventionality of papers and investigates their effects on academic impact to reveal the patterns of academic innovation. [Method/Process] Employing a Markov Chain Monte Carlo (MCMC)-based method, we measure the novelty and conventionality of 70,207 research papers in the library and information science field indexed in the Chinese Social Sciences Citation Index (CSSCI) during the 20-year period from 2000 to 2019, and analyze the effects of paper novelty and conventionality on disciplinary impact. [Results/Conclusion] Results indicate that, *ceteris paribus*, a one-unit increase in paper novelty increases the odds ratio of becoming a highly cited paper by 11%, whereas a one-unit increase in paper conventionality increases the odds ratio by 33%. Marginal effects analysis demonstrates that papers with simultaneously high novelty and high conventionality exhibit a higher likelihood of becoming highly cited compared to other paper types. Moreover, over time, the effect of novelty on the probability of papers becoming highly cited gradually weakens, while the effect of conventionality gradually strengthens. Simultaneously, author team size exerts a significant influence on paper novelty, and this influence intensifies over time. These findings underscore the characteristic of upholding tradition while fostering innovation in China' s LIS field, providing new empirical foundations for understanding academic innovation patterns in this domain. Additionally, this study proposes a novel Bayesian statistics-based methodology distinct from traditional informetrics.

Full Text

Can Tradition Foster Innovation? An Analysis Based on the Novelty and Conventionality of Papers in China's Library and Information Science Field

Liang Xingkun

Department of Information Management, Peking University, Beijing 100871

Abstract: [Purpose/Significance] This study examines papers in China's library and information science (LIS) field to measure their novelty and conventionality, explore how these attributes affect academic impact, and reveal the underlying patterns of scholarly innovation. [Method/Process] Using an optimized Markov Chain Monte Carlo (MCMC) method, we measured the novelty and conventionality of 70,207 research papers in library and information science from the Chinese Social Sciences Citation Index (CSSCI) between 2000 and 2019, and analyzed their effects on disciplinary impact. [Result/Conclusion] The results show that, holding other factors constant, a one-unit increase in paper novelty increases the odds of becoming a highly cited paper by 11% ($p < 0.0001$), while a one-unit increase in conventionality increases these odds by 33% ($p < 0.0001$). Marginal effect analysis reveals that papers with both high novelty and high conventionality have a higher probability of becoming highly cited than other types. Over time, the impact of novelty on the probability of becoming highly cited gradually weakens, while the impact of conventionality gradually strengthens. Meanwhile, author team size has a significant effect on paper novelty, and this effect increases over time. These findings highlight the characteristics of "innovation through tradition" in China's LIS field and provide new empirical evidence for understanding the patterns of academic innovation. This paper also proposes a novel Bayesian statistical method that differs from traditional informetrics approaches.

Keywords: novelty; conventionality; disciplinary impact; highly cited paper; citation count; Markov Chain Monte Carlo (MCMC); team authors; Bayesian statistics

Classification Number: G25

1 Research Background

Research on innovation and development in the library and information science (LIS) field has long focused on macro-level direction setting and 梳理 [1-8], while neglecting micro-level exploration of academic innovation mechanisms among researchers. From a micro-level perspective, although foreign studies can provide reference points, academic innovation mechanisms differ significantly across cultural backgrounds and disciplinary contexts [14-15, 18]. Systematic empirical research on this issue is crucial for understanding innovation development in LIS

but has been lacking. Foreign studies have shown that impactful innovations typically combine both novel and conventional knowledge [9-13]. Researchers can create novel knowledge to break through existing research, thereby achieving novelty [9, 11, 14-15]. Simultaneously, they can creatively recombine existing knowledge to better inherit and integrate previous research, thereby achieving conventionality [9-10, 12, 14, 16-17]. Studies indicate that specific research outcomes usually possess both novelty and conventionality, but in varying degrees, and different combinations of these attributes produce different disciplinary impacts [9, 14-16]. Therefore, a core question in exploring innovation patterns in LIS is how researchers leverage knowledge with both novel and conventional attributes to produce impactful innovations.

To systematically investigate this issue, this study examines 70,207 research papers in library and information science from the Chinese Social Sciences Citation Index (CSSCI) between 2000 and 2019. Using an optimized Markov Chain Monte Carlo (MCMC) method, we measure these papers' novelty and conventionality, analyze their effects on disciplinary impact and temporal trends, and further explore the micro-level mechanisms of novelty and conventionality formation from the perspective of author teams, providing empirical support for comprehensively understanding academic innovation mechanisms in China's LIS field.

2 Research Hypotheses

2.1 Relationship Between Novelty, Conventionality, and Academic Impact

Novelty is a crucial attribute of academic output. It is relative, referring to content that is rarer compared to certain standards [9, 14-15, 18]. These standards may be existing knowledge in a discipline [9-10] or a broader knowledge system of an era [18]. For example, using advanced algorithms for data analysis and mining is no longer novel in LIS but may remain highly novel in other disciplines. Highly novel research requires researchers to generate innovative or breakthrough ideas through creativity, such as new theories, methods, or topics [9, 14, 17]. These new ideas often inspire other researchers and generate stronger disciplinary impact [9, 11]. Uzzi et al.'s analysis of 17.9 million papers from various fields in Web of Science found that, after controlling for relevant factors, highly novel papers had significantly higher probabilities of becoming highly cited [9]. Subsequent studies on novelty in other types of English research have reached consistent conclusions [14-18].

This effect of novelty is also evident in China's LIS research. For instance, in the early development of Chinese library science theory, many influential theories were localized, formed by adapting and transcending Western library science theories in the Chinese context, demonstrating clear novelty [19]. During periods of information technology transformation, new concepts like digital libraries and

smart libraries attracted substantial attention and became research frontiers [20], while corresponding new services like knowledge services and smart services also became hot topics with widespread influence [21]. Based on this, we propose:

Hypothesis 1: In LIS, the stronger a paper's novelty, the greater its disciplinary impact.

Although novelty reflects creativity in research, many studies note that novelty is a double-edged sword for researchers [22]. On one hand, pursuing novelty increases uncertainty, potentially reducing productivity and extending research cycles [17, 23]. On the other hand, due to peer communication factors, highly novel results are harder to understand and recognize in the short term, leading to negative effects [15-25] such as lower publication rates [17], undervaluation of research contributions, and marginalization of authors [24]. Highly novel research proposals also face greater misunderstanding and lower scores in grant reviews [25]. This bias against novelty is not unique to research; it exists in other domains like business [26-28]. Some studies suggest this bias stems not only from bounded rationality in cost reduction and practicality but also from the mechanism of understanding new ideas—people use their experiences and knowledge to fill information gaps between new ideas and existing cognition, and highly novel content creates gaps that are difficult to bridge, hindering comprehension and acceptance [28].

Since novelty is a double-edged sword, researchers often maintain certain conventionality while pursuing novelty. “Conventionality” is not the opposite of novelty but rather reflects close association with existing knowledge and theoretical systems. Increased conventionality helps bridge the information gaps created by novelty, making novel content easier to accept [15, 28]. From another perspective, generating novel content in scientific research requires inheriting and developing existing knowledge or theories—through synthesizing existing knowledge, theories, techniques, or methods to form creative content. This characteristic is also evident in LIS research. For example, when Dewey discussed the necessity of professional education for librarians, he did not start from his newly created decimal classification system but from more familiar perspectives like librarian services to patrons and libraries' social education functions [29]. When Wang Chongmin elaborated on the bibliographic value of the *Siku Quanshu Zongmu*, he did not directly introduce its content features but started from the familiar historical background of the *Siku Quanshu* compilation, analyzing its cataloging methods based on the compilation process [30]. In summary, innovation in LIS cannot be separated from inheriting and developing existing knowledge [31]. Researchers must work on conventionality to demonstrate academic heritage and accumulation while pursuing novelty [9, 15]. Therefore, research with prominent conventionality helps other researchers understand development trajectories and generates broader impact. Based on this, we propose:

Hypothesis 2: In LIS, the stronger a paper's conventionality, the greater its disciplinary impact.

2.2 Relationship Between Team Authorship and Novelty/Conventionality

A temporal comparison reveals that the average number of authors per paper is increasing. As research complexity grows, most high-impact research today is completed by teams [9]. Studying team authorship helps clarify the micro-level mechanisms of novelty and conventionality formation. Many studies suggest that larger author teams create more diverse groups, integrating knowledge from different disciplines, depths, and breadths, thereby breaking research boundaries and producing more innovative and impactful results [9-14, 32-33]. Meanwhile, mutual selection mechanisms among team members help ensure quality floors, reducing the likelihood of poor outcomes [34]. However, larger teams also increase coordination costs and communication difficulties [32-33, 35], leading researchers to conclude that reasonably sized teams are more conducive to innovation. Characteristics of research collaboration in China's LIS field align with these findings, showing a trend of 2-3 person co-authorship and increasing high-impact papers by collaborative authors [36-38]. Therefore, we propose:

Hypothesis 3: In LIS, team-authored papers have greater disciplinary impact.

3 Methods and Data

3.1 Data Source and Preprocessing

Among various Chinese paper databases, we selected the CSSCI database because its data quality is superior to commercial databases like Wanfang, VIP, and CNKI, particularly in one-to-one correspondence between authors and affiliations, reducing disambiguation difficulty. We selected all papers published between 2000 and 2019 in library and information science journals indexed by CSSCI as our sample and performed data cleaning. First, we removed document types like announcements, editorials, and errata with limited research value, retaining 70,207 research papers and reviews. Next, we disambiguated authors and affiliations. Author disambiguation primarily used birthdate information from some papers' author biographies, combined with self-citations and co-authorship patterns. Affiliation information was not our focus, so we only processed institutional name changes at the university level using Ministry of Education announcements to ensure name consistency. Finally, we conducted manual verification through sampling to ensure disambiguation quality.

3.2 Variables and Measurement

3.2.1 Dependent Variable: Disciplinary Impact Many studies measure disciplinary impact using citation frequency or percentiles, with increasing emphasis on percentile advantages [39-42], but no widely accepted standard exists for defining highly cited papers. In practice, different thresholds like top 1%, 5%, or 10% are used [9-14, 17-18, 24]. Following relevant research [9-14] and

considering the sparser citation relationships in CSSCI, we selected the top 5% of citation counts as our standard for highly cited papers. We also used top 1% and top 10% as alternative measures to test statistical robustness. Additionally, because highly cited papers in LIS typically represent only a small fraction and cannot indicate general impact, we also used total citation counts as an alternative measure for further robustness testing.

3.2.2 Independent Variables: Novelty and Conventionality We employed an optimized MCMC method to measure novelty and conventionality [9, 14]. This approach treats novelty and conventionality as relative concepts that cannot be directly measured by observed frequencies but must be compared against a standard. A suitable and relatively objective approach is comparison with random probability [9, 14-15, 18], eliminating “common” or “uncommon” 偶然 factors from observations [9-14, 17] to demonstrate novelty. For example, if a concept appears 10 times in a discipline, whether it is “novel” or not cannot be determined directly from the observable “10 occurrences” but requires comparison with how often a randomly generated “new concept” might appear.

From this perspective, a paper’s references effectively reflect its disciplinary foundation and use of existing knowledge. The novelty (or conventionality) of research is embodied in the novelty (or conventionality) of co-cited references. Measuring the novelty of co-cited references yields the novelty of the research. The specific algorithm is:

1. Extract all pairwise journal combinations from a paper’s co-cited references.
2. Compile all pairwise journal combinations from all papers’ co-cited references, recording annual frequencies o .
3. Construct a citation network based on inter-paper citation relationships. To obtain comparable “random probabilities,” these studies use MCMC methods to randomize the observed citation network, simulating comparable networks. Since MCMC is computationally intensive, Uzzi et al. used Gibbs sampling to randomize the observed citation network [9], recommending exchange of citation relationships within each subgraph at least $n \times 2 \log_{10} n$ times, where n is the total number of citation relationships in subgraph i , to achieve MCMC convergence or approximate convergence. The resulting network maintains the same nodes, edges, in-degree, and out-degree as the observed network, with unchanged citation years and source papers. Two points require attention: (1) Gibbs sampling is not strictly random; (2) The recommended exchange frequency is not a necessary condition for MCMC convergence but an empirical rule for computationally intensive situations.
4. Repeat step (3) 10 times to obtain 10 randomized citation networks. For each network, count citation relationships per paper to derive co-citation frequencies for each journal pair, calculate mean e and standard deviation σ , then compute standardized scores (z) using:

$$z = \frac{o - e}{\sigma} \quad (1)$$

5. Based on the distribution of standardized scores z for a paper' s co-cited journal pairs, select appropriate percentiles as measures of novelty and conventionality. This method' s innovation lies in transforming a single observable frequency into multiple randomized networks, using standardized scores to eliminate random factors. Since mean e is based on random probability, combinations with $z < 0$ are rarer than random probability and thus novel, while those with $z > 0$ are more “conventional.”

[Figure 1: see original paper] shows the cumulative distribution of observed and simulated values for a randomly selected paper (Wu Weici and Gu Xiujie' s 2007 *Library Tribune* paper “2006 Academic Progress in Library Science”). The observed values are sparse, while simulated values are denser, providing a “baseline” for understanding relative novelty.

We optimized the novelty and conventionality measurement. Many studies emphasize that citation networks exhibit strong preferential attachment [43-46], where cited papers are more likely to receive further citations. Therefore, randomization should preserve this characteristic rather than assuming random citation processes. To better simulate this feature, we adapted the Chinese restaurant process from Dirichlet processes [47]:

1. Initially assume each paper has equal probability of being cited.
2. Using Gibbs sampling, when exchanging citation relationships between two observed papers, first check if the citation probability p exceeds a random value u between 0 and 1. Exchange only if $p > u$.
3. In subsequent steps, adjust u using the Dirichlet distribution' s probability density function based on observed citation patterns, making a paper' s citation probability increase with its citation count and its journal' s citation count.
4. Repeat steps (2) and (3) until all citation relationships are exchanged at least once. For our moderately sized network, we exchanged all relationships; for larger networks, Uzzi et al.' s empirical rule could be used.

[Figure 2: see original paper] illustrates the standardized score distribution for a sample paper (Ma Feicheng and Pei Lei' s 2005 *Journal of the China Society for Scientific and Technical Information* paper “Progress in Information Resource Sharing Practice and Theory”). The distribution is long-tailed before the 50th percentile and denser afterward, confirming that novel journal combinations are rarer.

To show overall patterns, we divided 2000-2019 into two decades (2000s and 2010s). [Figure 3: see original paper] shows novelty (10th percentile of standardized scores), and [Figure 4: see original paper] shows conventionality (50th percentile). Novelty increased over time: only 14.76% of papers had novelty < 0 in 2000-2009 ($n = 28,890$), rising to 28.83% in 2010-2019 ($n = 41,317$). Conven-

tionality remained high: 2.23% of papers had conventionality < 0 in 2000-2009, and 2.32% in 2010-2019. These results align with Uzzi et al. [9] but show much lower novelty and higher conventionality proportions.

3.2.3 Control Variables Following previous research, we controlled for publication year, author count, and journal. Publication year and journal are categorical variables with fixed effects controlled in models. Following Uzzi et al.'s team size classification and China's LIS collaboration patterns [9, 36], we categorized author count into: single author, dual authors, and team authors (3+).

3.3 Regression Model and Analysis Strategy

Corresponding to our hypotheses, we constructed the following regression model:

$$\text{Disciplinary Impact}_t = \beta_1 \times \text{Novelty}_t + \beta_2 \times \text{Conventionality}_t + \beta_3 \times \text{Author Count}_t + \beta_4 \times \text{Journal} + \beta_0 + \varepsilon \quad (2)$$

where β_0 is the intercept, ε is the error term, subscript t indicates publication year, and β_1 to β_4 are coefficients expected to be positive.

We used logistic regression to test how novelty, conventionality, and author count affect the probability of becoming highly cited. Note that when measuring novelty with standardized scores, smaller scores indicate greater novelty. To test whether greater novelty increases citation probability, we used the negative of the 10th percentile as our novelty measure, ensuring larger values indicate greater novelty. We strictly controlled for publication year and author count fixed effects.

4 Analysis Results

4.1 Relationship Between Novelty, Conventionality, and Disciplinary Impact

presents logistic regression results using top 5% citations as the highly cited threshold. Model 1 uses all papers, Model 2 uses 2000-2009 only, and Model 3 uses 2010-2019 only. Overall, paper novelty significantly increases the odds of becoming highly cited ($\beta = 0.109$, $p < 0.0001$), as does conventionality ($\beta = 0.292$, $p < 0.0001$). Comparing the two decades, novelty's impact weakened while conventionality's impact strengthened, providing strong empirical evidence for their roles.

Unlike linear regression, logistic regression relationships are nonlinear, and joint effects differ from coefficient sums [48]. To analyze joint effects, we used the 25th and 75th percentiles of novelty and conventionality as low/high values, creating

four combinations. [Figure 5: see original paper] shows that papers high in both attributes have a 7.08% probability of becoming highly cited, compared to 3.15% for those low in both, 4.90% for high novelty/low conventionality, and 4.60% for low novelty/high conventionality. Papers with high novelty have higher citation probabilities regardless of conventionality levels.

4.2 Relationship Between Author Count and Disciplinary Impact

Model 1 in shows dual-author papers have 43.17% higher odds of becoming highly cited than single-author papers ($p < 0.0001$), while team-author papers have 60.33% higher odds ($p < 0.0001$). Comparing Models 2 and 3 reveals that author count's impact strengthened from 2000-2019, demonstrating teams' growing advantage in disciplinary innovation.

[Figure 6: see original paper] shows mean standardized scores at various percentiles for different author counts across both decades. As author count increases, scores up to the 35th percentile generally decrease, with the trend becoming more pronounced over time. In 2000-2009, differences between single, dual, and team authors were small, but they widened in 2010-2019, particularly for novelty. This suggests that increased author count may promote novelty but has limited effect on conventionality.

4.3 Robustness of Results

Multicollinearity is low: Model 1's variance inflation factor is 2.63, with key independent variables below 1.7. White's test indicates heteroskedasticity ($\chi^2 = 1819.22$, $p < 0.001$), so we used robust standard errors [49].

We tested robustness using alternative measures. First, we used total citations within 4 years post-publication (based on LIS's 3-4 year citation half-life) in negative binomial regression (, Model 1), yielding consistent results. Second, we redefined highly cited as top 1% and top 10% (, Models 2-3), with consistent results except slightly reduced novelty significance in Model 2 ($p = 0.019$). Third, we used the 15th, 20th, 25th, 30th percentiles as novelty measures (, all showing significant effects ($p < 0.001$), confirming robustness.

5 Discussion

5.1 Measurement of Novelty and Conventionality

5.1.1 Comparison of Two Measurement Methods

Due to citation network characteristics, we adopted an optimized MCMC method. Comparing it with Uzzi et al.'s method [9] revealed advantages. We calculated standardized scores for all journal pairs using both methods (, finding no obvious differences in means, standard deviations, or extremes. However, correlation analysis yielded $r = 0.7975$ ($n = 2,433,255$, $p < 0.0001$).

To examine differences, we plotted both methods' results ([Figure 7: see original paper]). Paired t -tests showed equal variances, but our optimized method yielded significantly lower novelty means (10th percentile: $T = 2.22$, $p = 0.013$, one-tailed). The 15th and 20th percentiles were also lower ($T = 1.65$, $p = 0.049$; $T = 1.99$, $p = 0.023$), while no significant differences existed at higher percentiles. Correlations were weaker at lower percentiles (top 10%: $r = 0.6624$; top 30%: $r = 0.7031$; top 50%: $r = 0.7270$), indicating our method is more sensitive to novelty and better suited for sparse citation networks.

5.1.2 Applicability of Novelty and Conventionality Measurement To demonstrate applicability in China's LIS field, shows representative journal pairs from 2000 and 2019 with their standardized scores. The "accidental" appearance of these pairs in references aligns with expectations. For example, *Journal of Library Science in China* and *Journal of Academic Libraries* frequently co-occurred in both years. In 2000, the *Library Construction* and *Information Theory and Practice* pair was relatively novel. LIS-computer science combinations were novel around 2000 (*Journal of Intelligence* and *Acta Artificial Intelligence* had $z = 0$), but became conventional by 2019 (*Journal of the Association for Information Science and Technology* and *Journal of Software* had $z > 2$). Meanwhile, *Journal of Informetrics* and *Science* became strongly conventional, while *Journal of Intelligence* and *Social Sciences in China* remained highly novel. LIS-economics journal pairs show strong novelty: in 2000, *Journal of Library Science in China* and *Economic Research Journal*; in 2019, *Journal of Library Science in China* and *American Economic Review*.

Critiques of novelty measurement focus on whether it captures interdisciplinarity [50]. shows same-field journal pairs yield high standardized scores (high conventionality), while cross-field pairs don't necessarily yield high novelty. Even in 2019, non-LIS journal pairs had mean standardized scores > 0 , indicating they weren't novel compared to random probability. These findings align with Uzzi et al. [9], confirming our method's suitability.

To further distinguish novelty from interdisciplinarity, [Figure 8: see original paper] compares mean standardized scores for pure LIS journal pairs ($n_{2000} = 200$, $n_{2010} = 295$, $n_{2019} = 350$) and non-LIS pairs ($n_{2000} = 9,177$, $n_{2010} = 47,472$, $n_{2009} = 132,352$) across three years. [Figure 9: see original paper] shows mean standardized scores and 95% confidence intervals for CSSCI disciplinary categories. Journalism-communication journals have the most "common" cross-field combinations ($z = 0$), while literature, history, philosophy, and arts journals have lower means. Economics, sociology, psychology, natural resources, and geography have even lower means, with temporal variations further distinguishing novelty from interdisciplinarity.

5.2 "Innovation Through Tradition" in China's LIS Field

Results support Hypotheses 1 and 2, showing that both "tradition" (conventionality) and "innovation" (novelty) can produce impactful research. This differs

subtly from Kuhn's "paradigm" concept, where disciplinary change is "discontinuous" and dramatic, whereas LIS innovation occurs through inheritance. From Model 1, a one-unit increase in conventionality increases the odds of becoming highly cited by 33.86% ($p < 0.0001$), while a one-unit increase in novelty increases these odds by 11.48% ($p < 0.0001$). Interpretation requires caution: although conventionality's coefficient is larger, its mean is 2.632 (max = 3.015), making a one-unit increase unlikely. Novelty's mean is only 1.369277 (max = 3.015), making a one-unit increase plausible. Considering value ranges, both attributes' impacts are substantial. For example, novelty = -1.5 ranks in the top 8.18%, while novelty = -0.5 ranks in the top 22.39%. These results show that highly cited papers in China's LIS field combine both attributes, demonstrating "innovation through tradition."

Marginal effect analysis ([Figure 5: see original paper]) further clarifies this pattern. Papers high in both attributes are most likely to become highly cited, with 1.25 times higher probability than those low in both, 44.49% higher than high novelty/low conventionality papers, and 53.91% higher than low novelty/high conventionality papers. This confirms that innovation built on sufficient academic tradition— "innovation through tradition" —has greater potential impact [51-52].

Comparing Models 2 and 3 shows novelty's impact weakening while conventionality's strengthens over time. This may reflect rapid IT development and iteration: LIS research closely follows new technologies but hasn't fully integrated them with core disciplinary questions. Nevertheless, the trend shows that highly cited papers typically inherit and develop existing knowledge, with increasing emphasis on disciplinary heritage. This indicates that "tradition" must be maintained before "innovation" can occur, and this "tradition" is not conservatism but tradition oriented toward innovation [7, 19, 31].

These findings have implications for achieving innovation through tradition. Regarding "tradition," LIS's emphasis on academic heritage suggests we should further summarize, review, and reflect on existing theories and methods. Continuous refinement, development, and improvement of existing research can form general, comprehensive, and pioneering new theories and methods. For example, when Liu Guojun discussed establishing a unified classification system for China, he first systematically reviewed classification development, deeply analyzed existing systems' limitations, and identified problems the new system should solve [53]. Similarly, the "five elements" theory of library science, though developed for traditional environments, remains relevant today but requires updating through theoretical development and empirical testing to form foundational theory.

Regarding "innovation," LIS innovation must center on core disciplinary issues. Knowledge, methods, and information technologies from other fields provide fertile soil, but these "new" elements must be rooted in LIS core concerns to create impactful innovation. We should learn from other fields' mature theories and methods based on our discipline's needs to achieve breakthroughs. [Figure 8: see original paper] and [Figure 9: see original paper] suggest future innovation

could involve greater integration with previously underexplored fields like economics, sociology, psychology, natural resources, and geography, examining LIS phenomena from individual, group, societal, and environmental perspectives to develop and improve existing patterns.

[Figure 6: see original paper] reveals that team collaboration in LIS not only enhances novelty but also maintains high conventionality. Team authors play a crucial role in “innovation through tradition,” with their importance increasing over time. However, like a decade ago [36], collaboration remains dominated by small 2-3 person teams (43.96% of author groups), with teams of 3+ comprising only 6.24%. This suggests researchers haven’t fully recognized or leveraged teams’ potential in disciplinary innovation. Current institutional practices in LIS schools, which often require first or corresponding authorship for hiring and promotion, create barriers to multi-author collaboration. Overcoming these constraints to unleash teams’ innovative potential is an important future challenge.

5.3 Methodological Contributions

This study makes methodological contributions beyond exploring innovation patterns. It is among the first in Chinese LIS to use MCMC methods. Statistically, MCMC differs fundamentally from frequentist approaches in traditional informetrics. Frequentist methods assume citation distributions follow specific forms (e.g., power law) and estimate parameters through maximum likelihood. While power law distributions have been validated to varying degrees [42-43, 54], real-world data often involve complex distributions that violate these assumptions or represent mixtures, making likelihood functions intractable.

MCMC, a Bayesian method, does not require assuming a population likelihood function. Instead, it uses Bayesian inference (e.g., MCMC) to construct posterior distributions from empirical distributions for parameter estimation and simulation. This approach is particularly suitable for complex citation distributions with intractable likelihood functions. Bayesian methods often have better statistical power with small samples or high uncertainty and can perform inferences impossible with frequentist methods, effectively complementing and enriching traditional informetrics.

References

- [1] Wu Weici, Luo Zhiyong. New Trends in Library Science Research for the 21st Century[J]. *Journal of Library Science in China*, 2000(6): 3-6.
- [2] Lai Maosheng. A New Starting Point for Information Science in the 21st Century[J]. *Journal of the China Society for Scientific and Technical Information*, 2000(1): 81.
- [3] Lai Maosheng. How Can Information Management Discipline Development Break Through the Bottleneck Paradox?[J]. *Journal of Academic Library and*

Information Science, 2019, 37(5): 3-7.

[4] Wu Weici. Thoughts on Library Science Research and Discipline Development in the New Era—Speech at the 12th National Library Science Doctoral Forum in 2018[J]. Library Journal, 2019, 38(6): 24.

[5] Zhang Jiuzhen. LIS Discipline Construction Needs to Align with New Era Development[J]. Library and Information, 2020(6): 17-18.

[6] Ma Feicheng, Li Zhiyuan. Development Prospects of Library and Information Science in China Under the New Liberal Arts Background[J]. Journal of Library Science in China, 2020, 46(6): 4-15.

[7] Sun Jianjun, Li Yang, Pei Lei. Reflections on LIS Transformation in the “Digital-Intelligent” Era[J]. Library and Information Knowledge, 2020(3): 22-27.

[8] Ke Ping. New LIS—Development of the First-Level Discipline of Library, Information and Archives Management in New Liberal Arts Construction[J]. Information and Documentation Services, 2021, 42(1): 15-20.

[9] Uzzi B, Mukherjee S, Stringer M, et al. Atypical combinations and scientific impact[J]. Science, 2013, 342(6157): 468-472.

[10] Weitzman ML. Recombinant growth[J]. Quarterly Journal of Economics, 1998, 113(2): 331-360.

[11] McKnight SL. Unconventional wisdom[J]. Cell, 2009, 138(5): 817-819.

[12] Schilling MA, Green E. Recombinant search and breakthrough idea generation: an analysis of high-impact papers in the social sciences[J]. Research Policy, 2011, 40(10): 1321-1331.

[13] Arts S, Fleming L. Paradise of novelty—or loss of human capital? Exploring new fields and inventive output[J]. Organization Science, 2019, 29(6): 1074-1092.

[14] Mukherjee S, Uzzi B, Jones B, et al. A new method for identifying recombinations of existing knowledge associated with high-impact innovation[J]. Journal of Product Innovation Management, 2016, 33(2): 224-236.

[15] Foster JG, Rzhetsky A, Evans JA. Tradition and innovation in scientists’ research strategies[J]. American Sociological Review, 2015, 80(5): 875-908.

[16] Wagner CS, Whetsell TA, Mukherjee S. International research collaboration: novelty, conventionality, and atypicality in knowledge recombination[J]. Research Policy, 2019, 48(8): 1260-1270.

[17] Leahey E, Beckman CM, Stanko TL. Prominent but less productive: the impact of interdisciplinarity on scientists’ research[J]. Administrative Science Quarterly, 2017, 62(1): 105-139.

- [18] Wang J, Veugelers R, Stephan P. Bias against novelty in science: a cautionary tale for users of bibliometric indicators[J]. *Research Policy*, 2017, 46(8): 1416-1436.
- [19] Liu Ziheng. Re-discussing the Localization of Library Science[J]. *Library and Information*, 2005(4): 13-15.
- [20] Wang Shiwei. On the Three Characteristics of Smart Libraries[J]. *Journal of Library Science in China*, 2012, 38(6): 22-28.
- [21] Zhang Xiaolin. Toward Knowledge Services: Finding Growth Points for Library and Information Work in the New Century[J]. *Journal of Library Science in China*, 2000(5): 30-35.
- [22] Kaplan S, Vakili K. The double-edged sword of recombination in breakthrough innovation[J]. *Strategic Management Journal*, 2015, 36(10): 1435-1457.
- [23] Iaria A, Schwarz C, Waldinger F. Frontier knowledge and scientific production: evidence from the collapse of international science[J]. *Quarterly Journal of Economics*, 2018, 133(2): 927-991.
- [24] Hofstra B, Kulkarni VV, Munoz-Najar Galvez S, et al. The diversity-innovation paradox in science[J]. *Proceedings of the National Academy of Sciences*, 2020, 117(17): 9284-9291.
- [25] Boudreau KJ, Guinan EC, Lakhani KR, et al. Looking across and looking beyond the knowledge frontier: intellectual distance, novelty, and resource allocation in science[J]. *Management Science*, 2016, 62(10): 2765-2783.
- [26] Lu S, Bartol KM, Venkataramani V, et al. Pitching novel ideas to the boss: the interactive effects of employees' idea enactment and influence tactics on creativity assessment and implementation[J]. *Academy of Management Journal*, 2019, 62(2): 579-606.
- [27] Criscuolo P, Dahlander L, Grohsjean T, et al. Evaluating novelty: the role of panels in the selection of R&D projects[J]. *Academy of Management Journal*, 2017, 60(2): 433-460.
- [28] Cooney G, Gilbert DT, Wilson TD. The novelty penalty: why do people like talking about new experiences but hearing about old ones?[J]. *Psychological Science*, 2017, 28(3): 380-394.
- [29] Dewey M. The profession[J]. *American Library Journal*, 1876, 1(1): 5-6.
- [30] Wang Chongmin. On the *Siku Quanshu Zongmu*[J]. *Peking University Journal (Humanities)*, 1964(2): 63-78.
- [31] Wang Zizhou. Academic Innovation Must Start with Academic History Research[J]. *Library and Information Service*, 2007(3): 5.
- [32] Lee Y, Walsh JP, Wang J. Creativity in scientific teams: unpacking novelty and impact[J]. *Research Policy*, 2015, 44(3): 684-697.

- [33] Teodoridis F. Understanding team knowledge production: the interrelated roles of technology and expertise[J]. *Management Science*, 2018, 64(8): 3625-3648.
- [34] Singh J, Fleming L. Lone inventors as sources of breakthroughs: myth or reality?[J]. *Management Science*, 2010, 56(1): 41-56.
- [35] Leahey E. From sole investigator to team scientist: trends in the practice and study of research collaboration[J]. *Annual Review of Sociology*, 2016, 42(1): 81-100.
- [36] Wei Ruibin. Research on the Status Quo of Scientific Collaboration in China' s Library and Information Science—CSSCI Data from 1998-2004[J]. *Library and Information Service*, 2006(1): 41-43, 48.
- [37] Li Changling, Wei Xuqiu, Cui Bin, et al. Analysis of Collaboration Network Structure in China's Library and Information Science from 2004-2013[J]. *Journal of Intelligence*, 2015, 34(3): 119-124, 143.
- [38] Zhou Xiaoying, Dong Wei, Zhu Xiaomei, et al. Analysis of Characteristics of High-Impact Papers and Disciplinary Differences in Library and Information Science[J]. *Journal of Library Science in China*, 2012, 38(4): 99-109.
- [39] Bornmann L. How to analyze percentile citation impact data meaningfully in bibliometrics: the statistical analysis of distributions, percentile rank classes, and top-cited papers[J]. *Journal of the American Society for Information Science and Technology*, 2013, 64(3): 587-595.
- [40] Bornmann L, Leydesdorff L, Mutz R. The use of percentiles and percentile rank classes in the analysis of bibliometric data: opportunities and limits[J]. *Journal of Informetrics*, 2013, 7(1): 158-165.
- [41] Aksnes DW, Langfeldt L, Wouters P. Citations, citation indicators, and research quality: an overview of basic concepts and theories[J]. *SAGE Open*, 2019, 9(1): 215824401982957.
- [42] Radicchi F, Fortunato S, Castellano C. Universality of citation distributions: toward an objective measure of scientific impact[J]. *Proceedings of the National Academy of Sciences*, 2008, 105(45): 17268-17272.
- [43] Price DDS. A general theory of bibliometric and other cumulative advantage processes[J]. *Journal of the American Society for Information Science*, 1976, 27(5): 292-306.
- [44] Barabási AL, Albert R. Emergence of scaling in random networks[J]. *Science*, 1999, 286(5439): 509-512.
- [45] Smolinsky L, Lerchen A, McDaniel A. Testing theories of preferential attachment in random networks of citations[J]. *Journal of the Association for Information Science and Technology*, 2015, 66(10): 2132-2145.

- [46] Wang M, Yu G, Yu D. Measuring the preferential attachment mechanism in citation networks[J]. *Physica A*, 2008, 387(18): 4692-4698.
- [47] Teh YW, Jordan MI, Beal MJ, et al. Hierarchical Dirichlet processes[J]. *Journal of the American Statistical Association*, 2006, 101(476): 1566-1581.
- [48] Liang Xingkun. Regression Analysis in LIS Research: Basic Principles[J]. *Library and Information Knowledge*, 2021, 38(3): 154-164.
- [49] Liang Xingkun. Regression Analysis in LIS Research: Common Issues[J]. *Library and Information Knowledge*, 2021, 38(3): 165-176.
- [50] Fontana M, Iori M, Montobbio F, et al. New and atypical combinations: an assessment of novelty and interdisciplinarity[J]. *Research Policy*, 2020, 49(7): 104063.
- [51] Ye Jiyuan, Chen CM. Upholding and Expanding: Reflections on the Scientific Positioning of Library and Information Science Education in China and the US[J]. *Journal of Library Science in China*, 2007(2): 18-23.
- [52] Fan Bingsi. Confusion, Hardship, and Glory in Library Science Theory—Sixty Years of Library Science and *Journal of Library Science in China*[J]. *Journal of Library Science in China*, 2017, 43(1): 4-16.
- [53] Liu Guojun. The Development of Chinese Book Classification—With a Brief Introduction to Current Library Classification Systems in China[J]. *Library Science Communication*, 1981(2): 46-59.
- [54] Tian Shenghui, Su Linwei, Zhao Xing. Empirical Study on Power-Law Relationships Between Multi-Level Citation Sources and Total Citations[J]. *Journal of the China Society for Scientific and Technical Information*, 2015, 34(10): 1024-1030.

Author Profile: Liang Xingkun, Assistant Professor, PhD, Email: lxk@pku.edu.cn

Received: August 1, 2022 **Revised:** September 21, 2022 **Pages:** 148-161
Responsible Editor: Yi Fei

Citation: Liang Xingkun. Can Tradition Foster Innovation? An Analysis Based on the Novelty and Conventionality of Papers in China's Library and Information Science Field[J]. *Library and Information Service*, 2022, 66(20): 148-161.

Abstract (English): [Purpose/Significance] With papers in the field of library and information science in China, this study measures their novelty and conventionality and explores their effects on academic impact to reveal the law of scholarly innovation. [Method/Process] This paper adopted a Markov Chain Monte Carlo (MCMC) approach to measure the novelty and conventionality

of 70,207 research papers in library and information science from the Chinese Social Sciences Citation Index (CSSCI) between 2000 and 2019, and examined their impacts on papers' disciplinary impact using robust logistic regression. [Result/Conclusion] The results show that, ceteris paribus, a one-unit increase in paper novelty increases the odds ratio of becoming a highly cited paper by 11% ($p < 0.0001$), while a one-unit increase in conventionality increases the odds ratio by 33% ($p < 0.0001$). Marginal effect analysis shows that papers with high novelty and high conventionality are more likely to become highly cited papers than other types. Over time, the impact of novelty on the probability of becoming highly cited gradually weakens, while the impact of conventionality gradually strengthens. Meanwhile, author team size has a significant effect on paper novelty, and this effect increases over time. These findings highlight the characteristics of "innovation through tradition" in China's LIS field and provide novel empirical evidence to understand the law of scientific innovation. This paper also proposes an alternative Bayesian statistical method for informetrics.

Keywords: novelty; conventionality; scientific impact; hit paper; paper citation; MCMC; team authors; Bayesian statistics

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.