https://doi.org/10.51408/issi2025_093

Does Distance Still Matter? The Impact of Geographic Patterns in Scientific Knowledge Sourcing on Invention Value

Guiyan Ou[1], Chaocheng He[2], Jiang Wu[3]

[1]*Ouguiyan@whu.edu.cn*
Wuhan University, School of Information Management, Wuhan (China)

[2]*he_chaocheng@whu.edu.cn*
Wuhan University, School of Information Management, Wuhan (China)
Wuhan Unversity Shenzhen Research Institute, Shenzhen, Guangdong (China)

[3]*jiangw@whu.edu.cn*
Wuhan University, School of Information Management, Wuhan (China)

Abstract

While the "death of distance" hypothesis suggests that advanced information technologies have diminished the role of geography in knowledge flows, the question "Does distance still matter?" remains debated in scientific knowledge utilization. This study examines the influence of geographic factors on invention value by analyzing the spatial patterns of scientific knowledge sources cited in patents. Using a sample of 463,393 science-based patents granted by the USPTO, we investigate three geographic dimensions: geographical distance, external knowledge proportion, and geographical diversity. Our analysis demonstrates that the geographical distance of scientific knowledge sources negatively impacts both the technical and economic value of inventions. Furthermore, we find curvilinear (inverted U-shaped) relationships between invention technical value and both the proportion of external scientific knowledge and geographical diversity of knowledge sources. Notably, higher proportions of external scientific knowledge demonstrate a significant negative association with the economic value of patented inventions. These results challenge the "death of distance" hypothesis and demonstrate that geographical patterns remain crucial in scientific knowledge utilization. Our findings advance the understanding of geography's role in science-based innovation and offer important implications for strategic knowledge sourcing in the digital age.

Introduction

While technological advancement has become increasingly dependent on scientific discovery, the translation of scientific innovation into technological innovation remains a complex and uncertain endeavor. Early empirical evidence demonstrates that the diffusion of ideas, knowledge, and innovations is fundamentally constrained by geographical factors (von Graevenitz et al., 2022). And the geographical factors can largely shape the landscape of scientific as well as technological innovation. However, extant literature has paid limited attention to examining how the geographic patterns of scientific knowledge sources may influence the efficacy of transforming scientific discoveries into technological innovations.

The "death of distance" hypothesis, emerging from the rapid advancement of information and communication technologies, suggests that geographical constraints on knowledge flows have diminished in the digital age. This hypothesis posits that modern technological infrastructure has created a "placeless" paradigm of knowledge dissemination, theoretically enabling efficient access to globally distributed scientific knowledge (Abramo et al., 2020). However, recent empirical evidence challenges this notion, indicating that while digital technologies have reduced some geographic barriers, the influence of spatial distance on knowledge flows remains substantial (von Graevenitz et al., 2022). This persistent role of geography in knowledge transfer raises a fundamental question that motivates our research: Does geographic distance still matter in the transformation of scientific knowledge into valuable technological innovations?

Scientific knowledge, characterized by its codified nature and open standards, readily transcends geographic boundaries (Wang, 2024). This knowledge can be classified as local or external based on its spatial acquisition patterns across various geographical scales, from municipal to national boundaries. Overreliance on local knowledge sources, while convenient, risks cognitive lock-in and redundancy, potentially hindering innovation in increasingly complex systems (Hohberger and Wilden, 2022). Thus, understanding how varying intensities of external scientific knowledge input shape technological innovation outcomes becomes crucial.

Moreover, extant research demonstrates that geographical diversity in knowledge acquisition significantly influences innovation performance (Subramaniam and Yound, 2005). For instance, Belderbos et al. (2018) suggested that geographical diversification in Corporate Venture Capital (CVC) portfolios directly enhances firm-level innovation performance. Yet, the complexity of processing knowledge from diverse contexts imposes substantial coordination and integration costs (Lahiri, 2010; Singh, 2008). Consequently, the impact of geographical diversity in scientific knowledge sources on technological innovation outcomes merits rigorous investigation.

Technological inventions, representing the initial realization of novel concepts and principles, form the foundational basis of technical innovation. The value of these inventions serves as the critical metric for innovation success, fundamentally reflecting innovation effectiveness. This study examines science-based innovation by focusing on the geographical patterns of scientific knowledge sources. Specifically, we address three critical questions:

（1）How does the geographical distance of scientific knowledge sourcing affect invention value?

（2）What is the impact of external scientific knowledge proportion on invention value?

（3）How does the geographical diversity of scientific knowledge sources influence invention value?

Data and Variables

Data Sources

This study examines the science-technology linkages by analyzing patent citations to scientific literature. The empirical analysis draws on two primary data sources. First, we utilize patents granted by the United States Patent and Trademark Office (USPTO) from 2001 to 2010. Second, we employ the "Reliance on Science" (RoS) dataset (Marx and Fuegi, 2020), which contains 42,822,458 citation links between patents (both US and non-US) and scientific publications indexed in Microsoft Academic Graph (MAG). Our final sample consists of 463,393 USPTO granted patents that cite at least one MAG-indexed publication during the 2001-2010 period. These patents from 6,472,102 unique patent-to-publication citation pairs, linking to 1,407,439 distinct scientific publications in the MAG database.

Variables

We assessed invention value through two distinct dimensions: technological impact and economic value. We operationalized technological impact through the count of forward citations received within five years of patent grant (*Citation_5*) (Hsu et al., 2021). For economic value, we employed a binary measure of patent assignment (*Is_Assignment*) (Kwon,2020), coded as 1 if the patent was transferred and 0 otherwise.

We investigated three independent variables: geographical distance (*Geo_Dist*), the proportion of external scientific knowledge (*Exter_Know_Ratio*), and geographical diversity (*Geo_Dive*). The first independent variable—*Geo_Dist*—captures the spatial separation between scientific paper authors and patent inventors. Following Gao and Rai's (2023) approach, we employed the Haversine formula for *Geo_Dist* calculation, which is specified as:

$$Distance_{a,b} = 2r \cdot \arcsin\left(\sqrt{sin^2\left(\frac{\varphi_a - \varphi_b}{2}\right) + \cos(\varphi_a)\cos(\varphi_b)\, sin^2\left(\frac{\lambda_{a-} - \lambda_b}{2}\right)}\right)$$

where φ_a and φ_b denote the latitudes of a and b, respectively, while λ_a and λ_b represent their corresponding longitudes. The constant r represents Earth's radius, which is approximated as 6,371 kilometers.

Exter_Know_Ratio is measured as the ratio of cited scientific papers originating from countries different from the inventor's home country. This metric is calculated as follows:

$$Exter\_Know\_Ratio = \frac{Num\_Paper_{nonlocal}}{Num\_Paper}$$

Where Num_Paper represents the total number of scientific papers cited by the patent, and $Num\_Paper_{nonlocal}$ refers to the count of cited papers from countries other than the inventor's home country. *Geo_Dive* captures the spatial distribution of cited scientific papers in patents, which we quantify using the Blau index:

$$Geo\_diversity = 1 - \sum_{i=1}^{C} p_i^2$$

where C denotes the total number of countries from which cited papers originate, and p represents the proportion of papers from each country. The resulting *Geo_Dive* measure ranges from 0 to 1, with higher values indicating greater geographical dispersion of scientific knowledge sources cited in the patent.

Our analysis accounts for multiple factors that potentially influence invention value (Zhu et al., 2022; Büttner et al., 2022; Poege et al., 2019), including PCT patent, patent scope, technological maturity, technological diversity, technological originality, number of inventors, number of applicants, backward patent citations, scientific publication count, scientific maturity, and impact of scientific publications. Additionally, we include dummy variables controlling for applicant organization type, application year, and technological field.

Given the count nature of Citations_5 and the binary structure of Is_Assignment, we employ negative binomial regression and logistic regression models for estimation, respectively.

Results

Table 1 summarizes our regression analyses examining the impact of three key variables (*Geo_Dist, Exter_Know_Ratio and Geo_Dive*) on invention value. The empirical results from Models 1 and 4 demonstrate that *Geo_Dist* has significant negative effects on both *Citations_5* (β=-0.039, p<0.001) and *Is_Assignment* (β=-0.041, p<0.001). These findings suggest that greater geographical distances in scientofic knowledge absorption are associated with reduced forward citations and lower probability of commercial transactions for patents.

Models 2 and 5 demonstrate significant negative coefficients for both linear and quadratic terms of *Exter_Know_Ratio* in predicting *Citations_5* and *Is_Assignment*. While this pattern initially suggests potential inverted U-shaped relationships, subsequent graphical analyses yield divergent findings. Figure 1(a) confirms a robust inverted U-shaped relationship between *Exter_Know_Ratio* and *Citations_5*, indicating an optimal level of external knowledge integration for maximizing citation impact. However, Figure 1(b) reveals that the estimated inflection point (-1.636) lies outside the observed data range, invalidating the hypothesized curvilinear relationship for *Is_Assignment*. Instead, *Exter_Know_Ratio* exhibits a consistently negative linear relationship with assignment probability, suggesting that increased

external knowledge integration systematically reduces commercial transfer likelihood.

Model 3 demonstrates significant negative coefficients for both the linear and quadratic terms in the relationship between Geo_Dive and Citations_5. The empirical visualization in Figure 1(c) corroborates this inverted U-shaped relationship, suggesting that while moderate levels of geographical diversity enhance patent citation performance, excessive spatial distribution introduces coordination challenges that ultimately impede innovation. Contrary to expectations, we found no significant relationship between the Exter_Know_Ratio and Is_Assignment.

Table 1. The Impact of Knowledge Sourcing Geographical Patterns on Invention Value.

Variables	Citations_5			Is_Assignment		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Geo_Dist	-0.039***			-0.041***		
Exter_Know_Ratio		-0.096***			-0.052***	
Exter_Know_Ratio 2		-0.044***			-0.016***	
Geo_Dive			-0.020***			0.009
Geo_Dive2			-0.039***			-0.003
PCT	-0.299***	-0.259***	-0.297***	-0.122***	-0.108***	-0.132***
Ln (Patent_Scope)	0.248***	0.244***	0.248***	0.086***	0.087***	0.089***
Ln (Tech_Maturity)	-0.179***	-0.179***	-0.181***	-0.155***	-0.154***	-0.156***
Ln (Tech_Diversity)	-0.130***	-0.124***	-0.127***	-0.084***	-0.084***	-0.086***
Tech_Originality	-0.142***	-0.139***	-0.145***	0.074***	0.069***	0.069***

Ln (Num_Inventors)	0.080***	0.080***	0.081***	0.057***	0.054***	0.054***
Ln (Num_Applicants)	-0.103***	-0.089***	-0.099***	1.121***	1.119***	1.115***
Ln (Num_PR)	0.396***	0.390***	0.396***	0.150***	0.148***	0.151***
Ln (Num_NPR)	-0.019***	-0.032***	-0.008***	0.051***	0.043***	0.044***
Ln (Sci_Maturity)	-0.134***	-0.129***	-0.130***	-0.083***	-0.081***	-0.082***
Ln (Sci_Citation)	0.040***	0.036***	0.040***	0.001	0.002	0.004*
OrgTypes	Yes	Yes	Yes	Yes	Yes	Yes
Time	Yes	Yes	Yes	Yes	Yes	Yes
IPC	Yes	Yes	Yes	Yes	Yes	Yes
_cons	2.758***	2.828***	2.778***	0.182***	0.202***	0.187***
N	353205	369799	369799	353205	369799	369799

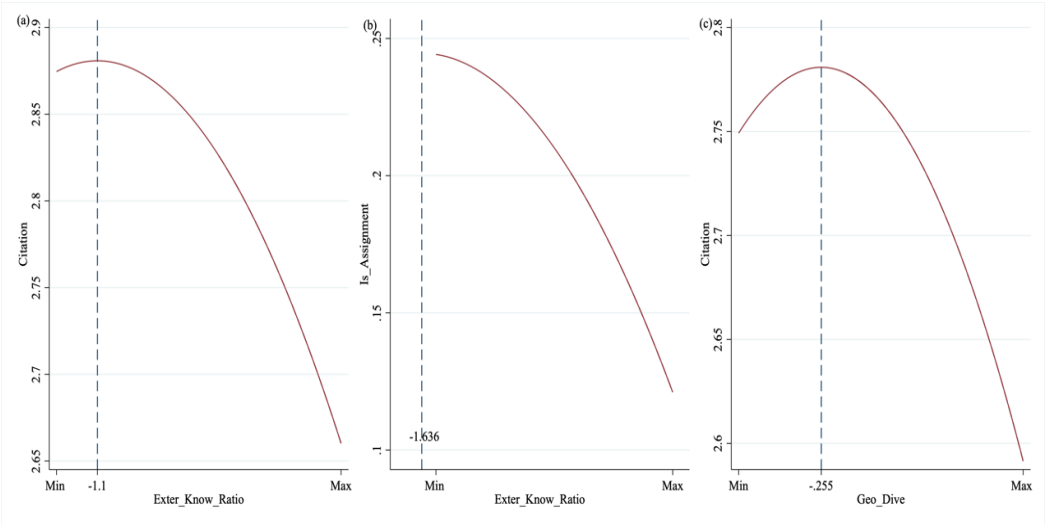

Figure 1. Nonlinear Effects of *Exter_Know_Ratio* and *Geo_Dive* on Invention Value with Identified Threshold Points (dashed lines).

Conclusion

This study advances our understanding of how geographical patterns of scientific knowledge sources influence invention value by examining three critical dimensions: geographical distance, external knowledge proportion, and geographical diversity. Through a systematic analysis of USPTO patent records integrated with RoS patent-science citation data, our empirical investigation yields several significant insights:

First, the geographical distance of scientific knowledge sources exhibits negative effects on both technological and economic value of patent inventions. While the advent of internet technologies has substantially reduced communication and interaction costs, potentially eliminating geographical barriers to knowledge dissemination in certain domains, our findings demonstrate that spatial distance remains a significant impediment in technological innovation. This persistent distance effect likely stems from the challenges in acquiring tacit knowledge necessary for effectively utilizing the codified knowledge embedded in scientific publications.

Second, we identify a nuanced relationship between external scientific knowledge proportion and invention value. Specifically, there exists an "optimal range" for external knowledge proportion in relation to technological value, while it demonstrates a consistently negative impact on economic value. Within this optimal range, increased input of external scientific knowledge correlates positively with patent citations. However, beyond a certain threshold, higher proportions of external scientific knowledge become detrimental to innovation quality. From a commercial perspective, greater reliance on external scientific knowledge in the technological development process corresponds to decreased likelihood of patent transfer.

Finally, our findings reveal an inverted U-shaped relationship between geographical diversity and technological value of patents, while showing no significant association with economic value. Patent forward citations increase with geographical diversity up to an optimal threshold, beyond which additional diversity becomes detrimental to technological impact.

This study has certain limitations. The temporal scope of our analysis (2001-2010) was deliberately chosen to accommodate patent transfer windows. Future research opportunities exist in extending this analysis to more recent periods by adopting economic value indicators that require shorter observation windows. Additionally, our analysis of patents across all technological domains does not account for varying degrees of science-dependency across different fields, potentially masking field-specific patterns.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (72204189), Guangdong Basic and Applied Basic Research Foundation (2022A1515110972) and Digital Intelligence Humanities Fundation of Wuhan University (2024SZWK023)

References

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2020). Does the geographic proximity effect on knowledge spillovers vary across research fields?. Scientometrics, 123, 1021-1036.

Belderbos, R., Jacob, J., & Lokshin, B. (2018). Corporate venture capital (CVC) investments and technological performance: Geographic diversity and the interplay with technology alliances. *Journal of Business Venturing*, *33*(1), 20-34.

Büttner, B., Firat, M., & Raiteri, E. (2022). Patents and knowledge diffusion: The impact of machine translation. *Research Policy*, *51*(10), 104584.

Gao, X., & Rai, V. (2023). Knowledge acquisition and innovation quality: The moderating role of geographical characteristics of technology. *Technovation*, *125*, 102766.

Hohberger, J., & Wilden, R. (2022). Geographic diversity of knowledge inputs: The importance of aligning locations of knowledge inputs and inventors. *Journal of Business Research*, *145*, 705-719.

Hsu, D. H., Hsu, P. H., Zhou, T., & Ziedonis, A. A. (2021). Benchmarking US university patent value and commercialization efforts: A new approach. *Research Policy*, *50*(1), 104076.

Kwon, S. (2020). How does patent transfer affect innovation of firms?. *Technological Forecasting and Social Change*, *154*, 119959.

Lahiri, N. (2010). Geographic distribution of R&D activity: how does it affect innovation quality?. *Academy of management journal*, *53*(5), 1194-1209.

Poege, F., Harhoff, D., Gaessler, F., & Baruffaldi, S. (2019). Science quality and the value of inventions. *Science advances*, *5*(12), eaay7323.

Singh, J. (2008). Distributed R&D, cross-regional knowledge integration and quality of innovative output. *Research Policy*, *37*(1), 77-96.

Subramaniam, M., & Youndt, M. A. (2005). The influence of intellectual capital on the types of innovative capabilities. *Academy of Management journal*, *48*(3), 450-463.

von Graevenitz, G., Graham, S. J., & Myers, A. F. (2022). Distance (still) hampers diffusion of innovations. *Regional Studies*, *56*(2), 227-241.

Wang, F. (2024). Does the recombination of distant scientific knowledge generate valuable inventions? An analysis of pharmaceutical patents. *Technovation, 130*, 102947.

Zhu, K., Malhotra, S., & Li, Y. (2022). Technological diversity of patent applications and decision pendency. *Research Policy*, *51*(1), 104364.