https://doi.org/10.51408/issi2025_186

Exploring institutional type composition in scientific collaboration and its role in scientific impact

Shuang Liang[1], Zhiyu Tao[2], Qingshan Zhou[3]

[1]liangs@stu.pku.edu.cn, [3]zqs@pku.edu.cn
Department of Information Management, Peking University, Beijing (China)

[2]taozhiyu@mail.las.ac.cn
National Science Library, Chinese Academy of Sciences, Beijing (China)
Department of Information Resources Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing (China)

Introduction

Many significant innovations and advancements in scientific research are achieved through inter-sector collaboration. With the continuous societal development, there is an increasing call for closer collaborative relationships among institutions, aiming to jointly address complex and evolving social challenges as well as technological issues. Interinstitutional collaboration integrates the unique resources and strengths of different institutions, playing a crucial role in improving research performance. It is a strategic approach for academic institutions to enhance funding acquisition and increase academic visibility (Zhou & Tian, 2014). Moreover, it also exerts a significant influence on researchers' academic performance, such as the number of publications and H-index (Bikard, Vakili & Teodoridis, 2019; Zhang & Wang, 2017).

Existing research has largely centered on universities, industries, and government institutions, often overlooking the roles and potential impacts of other institutional types, as well as the broader implications of institutional type composition in scientific innovation. In this study, we undertake a comprehensive examination of the roles of eight institutional types in scientific collaboration and investigate how institutional diversity and power structure influence research performance.

Methods

We obtained information of 26,998 institutions from Sciscinet (Lin et al., 2023). Institution type were further obtained from ROR (Research Organization Registry). Through records matching, all institutions were classified into eight categories: education, company, facility, government, healthcare, nonprofit, archive, and other. We extracted 8,454,850 records of multi-institutional collaborations from 1980 to 2021, each containing institutional information and five-year citation data (C5).

Institutional type diversity is defined as the number of different institutional types involved in academic collaboration. The formula is as follows:

$$D_i = \sum_{k=1}^{K} I_{ik}$$

where D_i denotes the institutional type diversity of paper i, K is the total number of institutional types, which is 8 in this study. If institutional type k is involved in paper i, then $I_{ik} = 1$, otherwise, $I_{ik} = 0$.

We use the Herfindahl Index (HI) to measure the institutional power centralization in academic collaboration:

$$P_i = \sum_{k=1}^{K} \left(\frac{ins\_num_{i,k}}{\sum_{j=1}^{K} ins\_num_{i,j}} \right)^2$$

where P_i denotes the institutional power centralization for paper i. $ins\_num_{i,k}$ represents the number of institutions of type k in paper i. K is the total number of institutional types, with a value of 8. A lower P_i indicates a more balanced distribution of institutional power. In contrast, papers with a high P_i have a higher degree of centralization in institutional power.

Results

Institutional type and scientific impact

We analyzed the number of publications and the 5-year citation performance for different institutional types (Fig.1). Nonprofit institutions demonstrate exceptionally excellent citation performance compared to other types. We created dummy variables for institutional types and controlled for other variables. Taking education institutions as the baseline, we performed a regression analysis to examine the effect of different institutional types on citations (Tab.1).

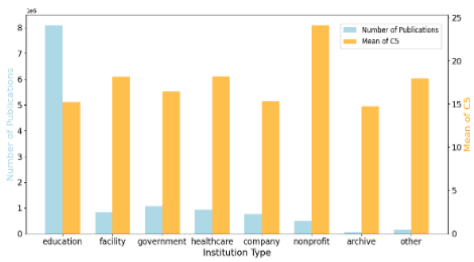

Figure 1. Publication count, C5 performance for different institutional types.

Table 1. Multivariable regression for institutional types and citation impact.

	logC5
company	0.0071** (0.002)
facility	0.1089*** (0.002)
healthcare	0.1119*** (0.004)
government	0.0835*** (0.002)
nonprofit	0.2280*** (0.003)
archive	-0.0469*** (0.005)
other	0.0921*** (0.003)
control	Yes
const	1.1639*** (0.011)
Obs	8454850
R^2	0.211

Note: *** p<0.001, ** p<0.01, * p<0.05, with robust standard errors in parentheses.

Institutional type diversity and scientific impact

We depicted the citation performance of different institutional type diversities (Fig.2) and found that collaboration across more institutional types contributes to higher citation counts. To further investigate whether there is a relationship between the institutional type diversity and citation impact of papers, we conducted the following multivariable regression analysis and reported the results in Table 2.

$$C_i = \alpha + \beta_1(D_i) + \beta_2(D_i^2) + \beta_3(Controls) + \varepsilon_i$$

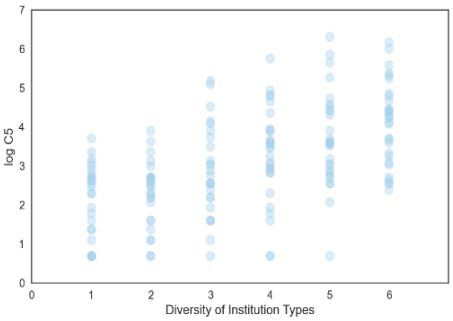

Figure 2. The relationship between institutional type diversity and citation impact.

Table 2. Multivariable regression for institutional type diversity and citation impact.

	logC5	*logC5*
type diversity	0.1168***(0.002)	0.0873***(0.004)
type diversity^2		0.0087***(0.001)
Control	Yes	Yes
const	1.0464***(0.009)	1.0714***(0.011)
Obs	8454840	8454840
R^2	0.210	0.210

Institutional power centralization and scientific impact

We further examined citation patterns across varying levels of institutional power centralization. The results reveal an overall decline in citation impact as power centralization increases (Fig.3). We conducted the following regression analysis to examine the potential nonlinear relationship between institutional power centralization and scientific impact (Tab.3). The estimated turning point occurs at 1.178, beyond the observed range of power centralization. Accordingly, citation impact consistently decreases with increasing centralization of institutional power.

$$C_i = \alpha + \beta_1(P_i) + \beta_2(P_i^2) + \beta_3(Controls) + \varepsilon_i$$

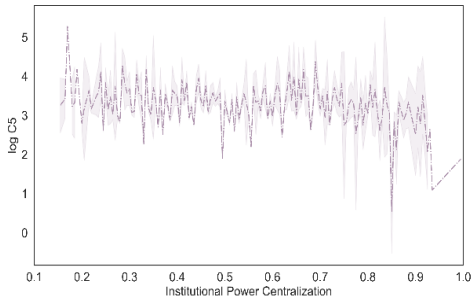

Figure 3. The relationship between institutional power centralization and citation impact.

Table 3. Multivariable regression for institutional power centralization and citation impact.

	logC5	*logC5*
power centralization	-0.2432***(0.004)	-0.6583***(0.028)
power centralization^2		0.2794***(0.018)
Control	Yes	Yes
const	1.3780***(0.014)	1.5139***(0.017)
Obs	8454850	8454850
R^2	0.209	0.209

Conclusion

This study offers a comprehensive categorization of institutional types and identifies statistically significant relationships between institutional type, institutional type diversity, institutional power centralization, and scientific impact. Greater institutional diversity is positively correlated with a significant increase in citation impact. Nevertheless, excessive power centralization in inter-sector institutional collaborations appears to hinder citation performance. The results provide valuable insights for research management and the development of institutional collaboration strategies.

References

Bikard, M., Vakili, K., & Teodoridis, F. (2019). When collaboration bridges institutions: The impact of university-industry collaboration on academic productivity. *Organization Science, 30*(2), 426-445.

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.

Zhang, B., & Wang, X. (2017). Empirical study on influence of university-industry collaboration on research performance and moderating effect of social capital: Evidence from engineering academics in China. *Scientometrics, 113*(1), 257-277.

Zhou, P., & Tian, H. (2014). Funded collaboration research in mathematics in China. *Scientometrics, 99*(3), 695-715.