STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at https://hdl.handle.net/1887/64521

ISBN: 978-90-9031204-0

23rd International Conference on Science and Technology Indicators (STI 2018)
"Science, Technology and Innovation indicators in transition"
12 - 14 September 2018 | Leiden, The Netherlands
#STI18LDN

Funding, collaboration and research performance: A comparative study of leading universities in China and the USA[1]

Ping Zhou*, and Xiaojing Cai*

* *pingzhou@zju.edu.cn; cai_xj@zju.edu.cn*
Department of Information Resources Management, Zhejiang University, No. 866 Yuhangtang Road, Hangzhou, 310058 (China)

Introduction

Investment in Research & development (R&D) is the guarantee of knowledge creation and technological innovation. Except for 2014 and 2015, China's investment in R&D had been in double-digit growth[2]. Investment in basic research grows even faster. Take the year 2016 for example, the growth rate of R&D was 10.6%, whereas that of basic research was 14.9%[3]. As the most important agency providing competitive grant in basic research, the National Natural Science Foundation of China (NSFC) had a budget increase of 15.2% in 2017[4]. The growth rate of R&D of the USA was 2.67% in 2016, and that of basic research was 3.38%[5]. The budget growth rate of the National Science Foundation (NSF) was higher (6.7%) in 2017[6].

The continuous growth of investment in R&D has caused widespread attention of stakeholders including researchers in science management and research evaluation. Since the time when Thomson Reuters (now Clarivate) started to index funding acknowledgement of journal publications indexed in the Science Citation Index (SCI) in 2008 and the Social Science Citation Index (SSCI) in 2015 (Alvarez-Bornstein, Morillo, & Bordons, 2017), studies analysing funding effect have been growing rapidly ((Paul-Hus, Desrochers, & Costas, 2016; Zhou & Tian, 2014; Wang, et al., 2012). Compared with unfunded publications, funded publications present higher citation impact (e.g., Costas & van Leeuwen, 2012; Jacob & Lefgren, 2011; Payne & Siow, 2003; Song, et al., 2016). Funding may also promote collaboration (Clark & Llorens, 2012; Chen, Huang, & Liang, 2010), and thus further rise citation impact (Sooryamoorthy, 2017; Wagner & Jonkers, 2017; Zhou & Glänzel, 2010). Funding support varies among countries and disciplines. China has the highest share of funded publications. With regard to funding support among disciplines (Costas & van Leeuwen, 2012), the humanities and social sciences receive far less support compared with the natural sciences (Tang, Hu, & Liu, 2017).

Different countries may have different arrangement in supporting competitive research. Take the first two publication producers, China and the USA, for example, China has only the

[1] This study receives funding from National Natural Science Foundation of China (grant number: 71473219).
[2] http://data.stats.gov.cn/easyquery.htm?cn=C01. Retrieved on 25th Oct, 2017.
[3] http://www.stats.gov.cn/tjsj/zxfb/201710/t20171009_1540386.html. Retrieved on 25th Oct, 2017.
[4] http://www.nsfc.gov.cn/publish/portal0/tab401/info68361.htm. Retrieved on 27th Oct, 2017.
[5] https://www.aaas.org/page/historical-trends-federal-rd#National. Retrieved on 27th Oct, 2017.
[6] https://www.nsf.gov/about/budget/fy2017/index.jsp. Retrieved on 27th Oct, 2017.

NSFC as the major grant source supporting competitive research in natural sciences, engineering and medical sciences. In the USA, however, responsibilities are dispersed across different agencies such as the NSF, National Institutes of Health (NIH), and so on (National Science Board, 2018). Funding agencies are also research targets of researchers. For example, Campbell and colleagues (2010) focused on the evaluation practice of National Cancer Institute of Canada (NCIC) and proved that bibliometric tools may help improve selection of outstanding researchers. Funding may promote establishment of international collaboration (Martinez et al., 2016).

In the development of science fields, funding agencies play important roles. For instance, both China's NSFC and US NSF have projects supporting big data research, but through different ways: NSFC projects are specialized in particular disciplines rather than being tagged as interdisciplinary or multidisciplinary, but for the NSF, the degree of interdisciplinarity is more pronounced. In other words, the NSF is more capable of leveraging funding in multiple areas, while NSFC is more direct to target resources (Huang et al., 2016). It is expected that more difference between the two agencies may exist. Clarifying characteristics of the two agencies in supporting competitive research may help understand the role of funding agencies with similar functions in supporting research and development, so as to provide reference to funding policy makers and researchers. The current study will analyze the effect of funding and collaboration in Chemistry research reflected by journal publications of domestic top-3 universities of China and the USA, with specific focus on research funded by the NSFC of China and the NSF of the USA.

Data and method

Publication data were downloaded from the SCI of Clarivate. The reason we choose Chemistry is that Chemistry is a traditional discipline in the natural sciences and is well covered by the SCI, and supporting basic research in the natural sciences is the common and basic responsibility of both NSFC and NSF. In view of the fact that leading universities are the main receivers of national funding, we select domestic top-3 universities of China and USA as the target sets. Based on the QS World University Rankings by Chemistry Subject[7], the top-3 in China are Tsinghua University, Peking University, Fudan University of China, and in the US are MIT, University of California Berkeley, and Stanford University.

Data

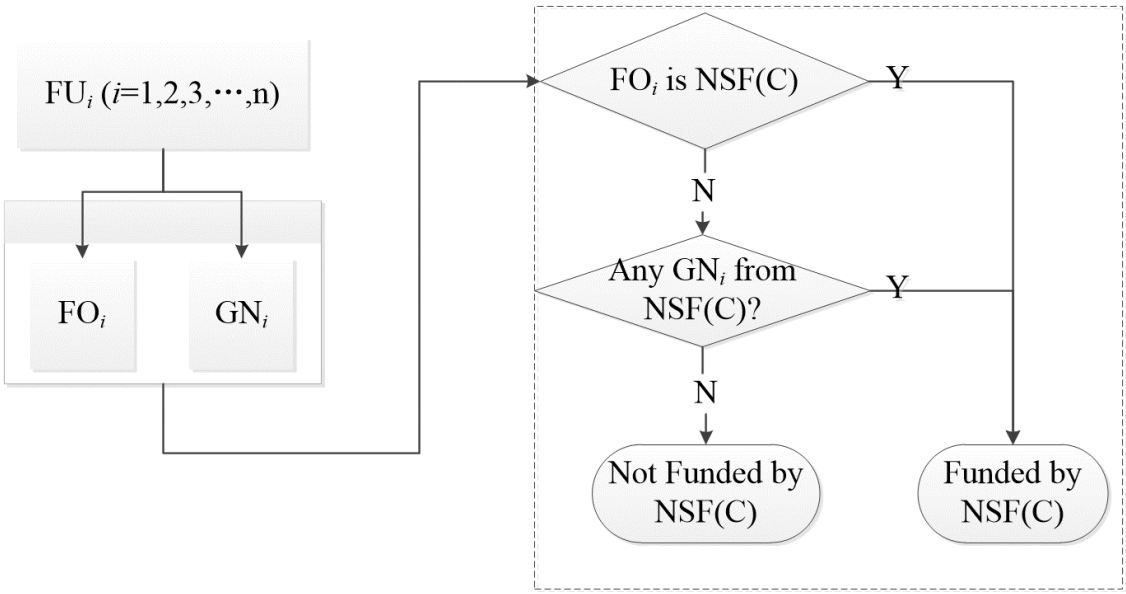
Publications of the targeted universities in Chemistry in 2009-2015 indexed in the SCI were retrieved using the institution names. Document types were limited to *Article* and *Review*. In total 31,022 papers, of which 18,808 from China and 12,394 from the USA were obtained[8]. Publications acknowledging funding support in the FU field are considered as funded publications, and thus the top-3 Chinese and US universities have respectively 17,158 and 11,417 funded publications in 2009-2015.

Funding information from the FU field of the SCI cannot be used directly due to a lot of mistakes that may be brought by authors and database builders (Wang et al., 2012). Authors may write the same funding agency in various ways, let alone wrong spellings. Database builders may make mistakes in extracting funding information from publication

[7] Data source: https://www.topuniversities.com/university-rankings/university-subject-rankings/2016/chemistry. Data retrieved on 16th July, 2017.
[8] Publication data retrieved in Nov, 2016.

acknowledgements. Identifying funded projects solely by project numbers is also not possible, because not every author provides such information. To ensure data of funding information of NSFC or NSF be as accurate as possible, we carry out data processing procedures as shown in Figure 1.

Figure 1: Process of identifying NSFC or NSF funded publications.



It is possible that more than one funding agencies are acknowledged. The Clarivate uses semicolon to separate different funding agencies in the same FU field. In such situation, text in the FU field is parsed into pieces by semicolon. Each piece is further split into two parts, namely, funding organization (FO) and grant number (GN) forming two datasets of FO and GN for further cleaning as shown in Fig 1. Based on acknowledgement of funding agencies in the publication dataset, we have collected a set of possible expressions for either NSFC or NSF to match with funding agency names in FO. We also use grant number identify funding agencies. In the end, respectively 14,265 and 4,097 publications are identified as funded by the NSFC or NSF.

Methods

The analysis is carried out from two perspectives – funding and collaboration. For exploration of funding effect, publications are classified into three types including *all publications, funded publications, and NSF(C)-funded publications*. For comparison of collaboration effect, publications are classified into types as shown in Table 1 according to Zhou and Tian (2014).

Table 1. Types of collaboration.

		Authors	**Institutions**	**Countries**
No collaboration (SingleAuth)		1		
Domestic collaboration	Within-institutional collaboration (WthinInst)	>1	1	
	Cross-institutional collaboration (CrossInst)	>1	>1	1
International collaboration (IntNatl)		>1	>1	>1

To explore the relationship between funding, collaboration and citation impact, we apply Negative Binomial (NB) regression model with robust standard errors. Stata 14.0 is used to conduct the regression analysis.

Results

The results are illustrated in two ways: descriptive and regression analysis. Due to limitation of word counts, here we just exhibit regression results. Conclusions are shown completely.

Descriptive analysis

In this section, we analyse development of publication productivity, citation impact of different types of publications defined by collaboration, funding, and NSFC or NSF funding of the top-3 universities of China and the USA.

Regression analysis

Negative Binomial (NB) regression analysis is done so as to illustrate relations between citation impact and different variables listed in Table 2. The regression is carried out in two steps. Step one is to analyse the effect of different collaboration types on citation impact of publications. Step two is to explore further on the effect of number of participants (i.e., authors, institutions and/or countries) in a specific collaboration type. Models 1, 2, and 3 in Table 3a respectively correspond to collaboration, funding, and funding plus collaboration. Models 4, 5, and 6 in Table 3b correspond to within-institutional, cross-institutional, and international collaboration respectively. Publication type being review, publication year, and journal rank (quartile in JCR) are controlled variables. Each of the four collaboration types (i.e., no collaboration, within-institutional collaboration, cross-institutional collaboration, and international collaboration) is set as a dummy variable. The first three variables are added to models 1 and 3, with international collaboration set as the baseline. The effect of these three types of collaboration in models 1 and 3 on the number of citations is tested by the test command (i.e., test $\alpha 1 = \alpha 2$, test $\alpha 1 = \alpha 3$, and test $\alpha 2 = \alpha 3$ in Table 3a).

Table 2. Variables used in NB regression analysis.

	Variable name	Description
Dependent variable	Citations	Number of citations received in three years including publishing year.
Independent variabls	Number of authors	Number of authors in a paper
	Number of institutions	Number of institutions in a paper
	Number of countries	Number of unique countries involved in a paper
	International collabiration	The value is set as 1 if two or more countries appear, or as 0 if only one country present in a paper.
	NSFC or NSF	The value is set as 1 if NSFC or NSF apprear, other wise as 0.
	Other Fundings	The value is set as 1 if a funding source other than NSFC or NSF appear, and as 0 if no funding appear.
Control variables	Review	The value is set as 1 for reviews and 0 for articles.
	Time length	The number of years between publishing year and 2018.
	Quartile	Quartile rank of journals in the publishing year.

Given that the quantity and relative share of uncollaborated publications in the total of either China or USA is quite low (respectively 85 or 0.45% and 184 or 1.48%), we will not analyze citation impact of uncollaborated publications so as to avoid influence of extremum. For publications of China, the citation counts of international collaboration are significantly higher than that of within-institutional or cross-institutional collaboration ($p < 0.01$), even though such types of collaboration are funded in general. There is no significant citation difference between publications of within- and institutional collaboration (Test $\alpha_2=\alpha_3$, $p > 0.1$). The positive effect of the NSFC in rising citation impact is obvious. The situation is somewhat different to publications of the USA, where no significant variation exists among three types of collaboration. In other words, citation impact of US publications is independent of collaboration types. Funding support, especially NSF funding, however, has positive effect.

Table 3a. Effect of collaboration and funding on Citation impact (Negative Binomial regression).

	China			USA		
	1a	2a	3a	1b	2b	3b
No Collaboration (α1)	-0.4118*** (0.1314)		-0.3617*** (0.1347)	-0.4131*** (0.1073)		-0.345*** (0.1053)
Within-institutional collaboration (α2)	-0.1919*** (0.0273)		-0.2191*** (0.0279)	0.0435 (0.029)		0.0349 (0.0287)
Cross-institutional collaboration (α3)	-0.1806*** (0.0261)		-0.224*** (0.0267)	0.0228 (0.0289)		0.0029 (0.029)
NSFC/NSF		0.162*** (0.0241)	0.2251*** (0.0242)		0.1923*** (0.0267)	0.1896*** (0.0265)
OtherFunding		0.3256*** (0.0264)	0.292*** (0.0264)		0.2651*** (0.0564)	0.2536*** (0.0563)
Review	Controlled					
Time length	Controlled					
Quartile	Controlled					
Const	1.6301*** (0.0804)	1.1126*** (0.0766)	1.2710*** (0.0788)	2.5123*** (0.1315)	2.2356*** (0.1408)	2.2448*** (0.1406)
Test α2=α3	--		--	--		--
Log pseudolikelihood	-64503	-64375	-64304	-47970	-47899	-47888
Wald chi2	5921.13***	5991.18***	6193.93***	1419.22***	1550.02***	1581.55***
Pseudo R2	0.0592	0.0515	0.0526	0.0231	0.0246	0.0248
Num. Objects	18,740	18,740	18,740	12,353	12,353	12,353
* p<0.1; ** p < 0.05; *** p< 0.01, -- p>0.1, robust standard errors in parentheses						

Both China and the USA share common features: Collaboration scale measured by the number of authors generates significantly positive effect on citation impact of any type of collaboration (i.e., within-institutional, cross-institutional, and international collaboration), effect of the number of institutions is, however, significant and negative. Increasing number of countries in international collaboration does not significantly generate effect on citation impact of Chinese publications, and may even have negative impact on US publications, although such effect is not significant (Table 3b).

To Chinese publications, the NSFC funding has positive and significant effect on citation impact of publications of within- and cross-institutional collaboration, but it is not the case to international collaboration. Other non-NSFC funding, however, generates positive and significant effect on all three types of collaboration. To US publications, NSF funding has significant and positive effect on within- and international collaboration, but the effect on cross-institutional collaboration is weak. Other non-NSFC funding has significant and positive effect on cross- and international collaboration, but with insignificant effect on within-institutional collaboration (Table 3b).

Table 3b. Effect of collaboration scale and funding on Citation impact (Negative Binomial regression).

	China			USA		
	4a WithinInst	5a CrossInst	6a IntNatl	4b WithinInst	5b CrossInst	6b IntNatl
Number of authors	0.0526***	0.0472***	0.0765***	0.0626***	0.0538***	0.0789***
	(0.0088)	(0.0061)	(0.008)	(0.0129)	(0.0082)	(0.0075)
Number of institutions		-0.0707***	-0.0691***		-0.0373*	-0.0775***
		(0.0221)	(0.0228)		(0.0227)	(0.0148)
Number of countries			0.0706			-0.0011
			(0.0447)			(0.028)
NSFC/NSF	0.2656***	0.3397***	0.0084	0.137***	0.0763*	0.2487***
	(0.0434)	(0.0379)	(0.0414)	(0.0423)	(0.0431)	(0.0419)
OtherFunding	0.2376***	0.2948***	0.1597**	0.1071	0.3011***	0.2476***
	(0.0413)	(0.038)	(0.0765)	(0.0925)	(0.0821)	(0.0819)
Review	Controlled					
Time length	Controlled					
Quartile	Controlled					
Const	0.7132***	0.7020***	1.7812***	2.3048***	2.1273***	1.6620***
	(0.1239)	(0.1358)	(0.2334)	(0.1394)	(0.2575)	(0.2261)
Log pseudolikelihood	-21892	-27360	-14544	-15665	-12538	-18607
Wald chi2	2591.57***	2906.29***	1037.35***	316.20***	631.66***	946.37***
Pseudo R2	0.0592	0.0586	0.0381	0.0214	0.0298	0.0353
Num. Objects	6,480	8,256	3,917	3,979	3,246	4,897
* p<0.1; ** p < 0.05; *** p< 0.01, robust standard errors in parentheses						

Conclusion and discussions

Publication production of the top-3 Chinese and US universities have kept linear growth, with the former more productive, leaving the latter far behind. Most publications of the two target groups are funded. NSFC funding contributes significantly more to Chinese publications than contribution of NSF to US publications. In contrast to the growing role of NSFC in publication production, that of NSF is decreasing slowly.

Funding support (including NSF or NSFC) has positive effect on citation impact. The effect of NSF is most prominent, whereas that of the NSFC is marginal. Citation impact of US publications, regardless of being funded or not, is significantly higher than that of Chinese publications. Almost all publications of the two target sets are collaborated. Cross-institutional collaboration takes the majority in China, whereas international collaboration takes the lead in the USA. Neither the NSFC nor NSF is active in supporting international collaboration.

With regard to the effect of collaboration types on citation impact, significant difference exists between the two countries. Generally speaking, international collaboration significantly improves citation impact of Chinese publications but has negative effect on US publications. Cross-institutional collaboration has negative effect on citation impact of publications of both countries. Similar conclusions apply to collaborated and funded publications, except for publications acknowledging NSF grant: international collaboration improves citation impact. The opposite results of international collaboration to the total versus NSF-funded publications of the top-3 US universities might be because of different roles of US scholars in international collaboration. In total publications, most international collaboration probably happened between US scholars and junior foreign scholars (e.g., visiting scholars, PhD or postdoctoral fellows), whereas international collaboration supported by the NSF requires that both US scholars and foreign scholars have established academic positions. Thus, it is not difficult to understand the different citation impact of international collaboration of the two types of publications.

Negative Binomial analysis confirms the effect of international collaboration and NSFC funding to citation impact of China. There is no significant citation difference between publications of within- and institutional collaboration. The situation is somewhat different to publications of the USA. Citation impact of US publications is independent of collaboration types. Funding support especially NSF funding has positive effect. Collaboration scale measured by the number of authors generates significantly positive effect on citation impact of any type of collaboration (i.e., within-institutional, cross-institutional, and international collaboration), effect of the number of institutions is, however, significant and negative. Increasing number of countries in international collaboration does not generate significant effect on citation impact of Chinese publications, and may result in negative impact to US publications, although such effect is not significant. To Chinese publications, the NSFC funding has positive and significant effect on citation impact of publications of within- and cross-institutional collaboration, but it is not the case to international collaboration. Other non-NSFC funding, however, generates positive and significant effect on all three types of collaboration. To citation impact of US publications, NSF funding has significant and positive effect on within- and international collaboration, but the effect on cross-institutional collaboration is weak. Other non-NSFC funding has significant and positive effect on cross- and international collaboration, but with insignificant effect on within-institutional collaboration.

References

Alvarez-Bornstein, B., Morillo, F., & Bordons, M. (2017). Funding acknowledgments in the Web of Science: completeness and accuracy of collected data. *Scientometrics*, *112*, 1793–1812.

Campbell, D., Picard-Aitken, M., Côté, G., Caruso, J., Valentim, R., Edmonds, S., Williams, G. T., et al. (2010). Bibliometrics as a Performance Measurement Tool for Research Evaluation: The Case of Research Funded by the National Cancer. *American Journal of Evaluation*, *31*(1), 66–83.

Chen, Y., Huang, B., & Liang, W. (2010). International cooperation in chemical sciences-the overview and perspectives. *Bulletin of National Natural Science Foundation of China*, (4), 211–215.

Clark, B. Y., & Llorens, J. J. (2012). Investments in Scientific Research: Examining the Funding Threshold Effects on Scientific Collaboration and Variation by Academic Discipline. *Policy Studies Journal*, *40*(4), 698–729.

Costas, R., & Van Leeuwen, T. N. (2012). Approaching the "reward triangle": General analysis of the presence of funding acknowledgments and "peer interactive communication" in scientific publications. *Journal of the American Society for Information Science and Technology*, *63*(8), 1647–1661.

Huang, Y., Zhang, Y., Youtie, J., Porter, A. L., & Wang, X. F. (2016). How does national scientific funding support emerging interdisciplinary research: A comparison study of big data research in the US and China. *PLoS ONE*, *11*(5), e0154509.

Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. *Journal of Public Economics*, *95*(9), 1168–1177.

Martinez, A., Epstein, C. S., & Parsad, A. (2016). Developing internationally engaged scientists and engineers: The effectiveness of an international postdoctoral fellowship program. *Research Evaluation*, *25*(2), 184–195.

National Science Board. (2018). Research and Development: U. S. Trends and International Comparisons. *Science & Engineering Indicators 2018*, 4.

Paul-Hus, A., Desrochers, N., & Costas, R. (2016). Characterization, description, and considerations for the use of funding acknowledgement data in Web of Science. *Scientometrics*, *108*(1), 167–182.

Payne, A. A., & Siow, A. (2003). Does federal research funding increase university research output? *Advances in Economic Analysis and Policy*, *31*(1), 1–24.

Song, Z., Guo, Y., & Li, D. (2016). Does science funding support increase research output-An empirical investigation based on propensity score stratification method. *Studies in Science of Science*, *34*(1), 116–121+160.

Sooryamoorthy, R. (2017). Do types of collaboration change citation? A scientometric analysis of social science publications in South Africa. *Scientometrics*, *111*, 379–400.

Tang, L., Hu, G., & Liu, W. (2017). Funding Acknowledgment Analysis: Queries and Caveats. *Journal Of The Association For Information Science And Technology*, *68*(3), 790–794.

Wagner, C. S., & Jonkers, K. (2017). Open countries have strong science. *Nature*, *550*, 32.

Wang, X., Liu, D., Ding, K., & Wang, X. (2012). Science funding and research output: A study on 10 countries. *Scientometrics*, *91*(2), 591–599.

Zhou, P., & Glänzel, W. (2010). In-depth analysis on China's international cooperation in science. *Scientometrics*, *82*(3), 597–612.

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