The MIT Press

RESEARCH ARTICLE

The dominance of big teams in China's scientific output

Linlin Liu[1] iD, Jianfei Yu[1] iD, Junming Huang[2,3] iD, Feng Xia[4] iD, and Tao Jia[1,5] iD

[1]College of Computer and Information Science, Southwest University, Chongqing, 400715, P. R. China
[2]Paul and Marcia Wythes Center on Contemporary China,
Princeton Institute for International and Regional Studies, Princeton University, Princeton, NJ 08540, USA
[3]Center for Complex Network Research, Northeastern University, Boston, Massachusetts 02115, USA
[4]School of Science, Engineering and Information Technology, Federation University Australia, Ballarat, VIC 3353, Australia
[5]Deakin-SWU Joint Research Center on Big Data, Southwest University, Chongqing, 400715, P. R. China

ABSTRACT

Modern science is dominated by scientific productions from teams. A recent finding shows that teams of both large and small sizes are essential in research, prompting us to analyze the extent to which a country's scientific work is carried out by big or small teams. Here, using over 26 million publications from Web of Science, we find that China's research output is more dominated by big teams than the rest of the world, which is particularly the case in fields of natural science. Despite the global trend that more papers are written by big teams, China's drop in small team output is much steeper. As teams in China shift from small to large size, the team diversity that is essential for innovative work does not increase as much as that in other countries. Using the national average as the baseline, we find that the National Natural Science Foundation of China (NSFC) supports fewer small teams than the National Science Foundation (NSF) of the United States does, implying that big teams are preferred by grant agencies in China. Our finding provides new insights into the concern of originality and innovation in China, which indicates a need to balance small and big teams.

1. INTRODUCTION

Modern science has witnessed the increasing dominance of teams. Single-author papers, though not yet as distinctly as what Price predicted in 1963 (Price, 1963), have undergone a sharp drop, taking only a small portion of all publications (Barlow, Stephens, et al., 2018; Larivière, Gingras, et al., 2015; Wuchty, Jones, & Uzzi, 2007). Teams become the driving force of science because not only is the problem to tackle more complex, but also the knowledge required is broader, which inevitably makes scientists more specialized (Jones, 2009; Leahey, 2016). Improvements of communication technology, the convenience of transportation and globalization also facilitate scientific collaborations. All of these make teams not only flourish but also grow in size (Gazni, Sugimoto, & Didegah, 2012; Larivière et al., 2015; Newman, 2001; Wu, Wang, & Evans, 2019). The average number of authors per publication increases every year and large teams involving more than 1,000 members have become common. In a recent paper studying the mass of the Higgs boson, the team size reached a record high of over 5,000 scientists (Castelvecchi, 2015).

The large team has clear advantages over the small team in solving complicated problems, securing research grants, receiving more citations on average, and publishing hit papers that

top the citation rankings (Cummings & Kiesler, 2007; Thelwall, 2019; Wuchty et al., 2007). Recent research shows, however, that bigger is not always better (Wu et al., 2019). Instead, small and large teams have distinct yet equally essential roles in science. Large teams tend to work in established fields and exploit existing problems. In contrast, small teams are better at exploring the frontier of science, generating new ideas, and opening up new problems that can disrupt science. To better promote science, a balance between small and large teams is needed (Azoulay, 2019), giving rise to an interesting question: To what extent is the research work of a nation carried out by big and small teams?

The answer to this question may have important implications for the scientific performance of a nation if we accept the fact that patterns observed in small and large teams are universal. While it is hard to argue whether a balance or an optimum is reached in a nation, it is still meaningful to compare the small vs. large team composition in different countries. This is of particular importance to China in the context of its long-term goal to be a global innovator (Phillips, 2016; Zhou, Lazonick, & Sun, 2016). Indeed, while China has grown to be the world's top scientific paper producer and citation receiver, it is often a worry that China's scientific work is weak in originality and innovation (Guo, Liu, et al., 2019; Huang, 2018; Xie, Zhang, & Lai, 2014). In this paper, we perform quantitative analyses on over 26 million papers published from 2000 to 2017. We find that China is indeed different from other countries. The percentage of China's scientific annual output from small teams is now the lowest in the world, after a sharp drop since 2000. As research teams in China shift from small to large size, the team diversity that is essential for innovative work has not increased as much as in other countries. Most work by big teams is still carried out in one or two institutes. The dominance of big teams in China may not be explained by the citation boost from the team size. While the number of citations on average increases with team size, the rate of increase is roughly the same in every country. However, the preference of funding agencies may be related to the lack of small teams in China. If small teams are more apt to perform disruptive research, the science community in China should be alerted, given the different statistics that China demonstrates.

2. DATA AND METHODS

2.1. Data Set

We use the publication data of the Web of Science (WoS), covering the Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI), and Arts & Humanities Citation Index (A & HCI) databases. There are over 26 million publications from year 2000 to 2017, including 18,295,191 articles, 3,646,465 meeting abstracts, 1,255,019 proceedings papers, 1,055,520 reviews, 970,649 editorial materials, 600,187 letters, 447,620 book reviews, 79,121 corrections, 32,205 biographical items, 29,115 news items, and more. The variety of document types naturally prompts us to check whether the conclusion would change when a different set of documents are considered. In particular, we perform a separate analysis by considering more "traditional" form of scientific papers, including article, review, letter, and conference proceeding. We find that there are only small changes to the statistics and the conclusions we draw remain the same.

2.2. Statistical Test

The statistical test is crucial when the sample size is relatively small. The *p*-value is usually reported to gauge whether the difference between the two measures is statistically significant. However, given the size of the data applied in our study, we find that most of our comparisons are statistically significant ($p \leq 0.05$). This can be demonstrated in a theoretical manner.

Consider a very general case in our analysis: There are two proportions p_1 and p_2 based on two samples with size n_1 and n_2, respectively. To test if p_1 is significantly bigger or smaller than p_2, we need to use the one-tail z-test. To simplify the model a bit more, let's consider the smaller one in p_1 and p_2 is p_o and the larger one is $p_o + \delta$. The two samples can be approximated with equal size that $n_1 = n_2 = n$. We can then plug the parameter p_o, δ and n into the calculation of the z-score (which consequently gives the p-value). When $n = 15,000$, the difference $\delta = 0.01$ is guaranteed to be statistically significant regardless of the p_o. Because most samples in this study have size greater than 15,000, it means that almost any virtual difference in the figure is statistically significant. Therefore, we choose not to report the p-value repetitively. If not otherwise mentioned, two proportions are statistically different. Indeed, there is only one instance (which is explicitly mentioned) in the paper, where the two measures are so close that the difference is not significant.

2.3. Country Allocation

We use straight counting by first affiliation (Huang, Lin, & Chen, 2011; Waltman & van Eck, 2015; Zheng, Zhao, et al., 2014). The country of a paper's first affiliation determines the country this paper belongs to. Other methods, such as whole counting and fractional counting (Kao, 2009; Larsen, 2008; Lewison, Purushotham, et al., 2010; Lin, Huang, & Chen, 2013; Sivertsen, Rousseau, & Zhang, 2019), are also widely applied to count publication numbers, but they may raise the issue of multiple counting, which can be a problem in the analysis. Previous work suggests that straight counting might be better when studying the scientific output at the country level (Huang et al., 2011). Another strategy of straight counting is to use the corresponding affiliation or the so-called "reprint address" in the WoS database (González-Alcaide, Park, et al., 2017; Kahn & MacGarvie, 2016; Mazloumian, Helbing, et al., 2013). We find that for more than 95% of papers, the reprint address and the first affiliation point to the same country. For simplicity and the ease of future reproduction of our analyses, we choose to use the first affiliation, as the information about corresponding affiliation may not be directly available in other databases. Finally, to eliminate possible bias caused by straight counting in dealing with papers by international collaborations, we also separately analyze publications by authors from the same country. We find that our conclusion is not affected (Supplementary Note 1).

2.4. Countries Considered

We include 15 countries in our analyses, which are roughly the top 15 countries for total scientific publications in our analysis (except for Turkey which ranks 16th). They are United States (US), China (CN), United Kingdom (GB), Germany (DE), Japan (JP), Italy (IT), France (FR), Canada (CA), India (IN), Korea (KR), Spain (ES), Australia (AU), Brazil (BR), Netherlands (NL), and Turkey (TR). Given China's huge annual production of scientific papers, it is less meaningful to compare it with countries of lower scientific output. Following typical practices, we use the scientific production from the mainland of China, Hong Kong, and Macau. We remove China when taking a global average to show a more vivid comparison between China and the rest of the world.

2.5. Big Teams and Small Teams

The terms *big team* and *small team* are relatively new and there is no defined hard cutoff between them. Previous work (Wu et al., 2019) considers team size (m) of no more than 3 or 4 members as small. In this work, we analyze all situations ($m \leq 3$, $m \leq 4$, $m \leq 5$) and find that our

conclusion in general is not affected by the choice of parameters. The only inconsistency is that $P(m \leq 5)$ of China is slightly higher than that of Japan and Italy, making China not the lowest, but the third lowest among the 15 countries. The value is still way below the global average. We present results based on $m \leq 4$ in the main text of the paper. The corresponding results for $m \leq 3$ and $m \leq 5$ are presented in the Supplementary Information.

2.6. Research Field of a Paper

WoS has approximately 250 subject areas characterizing different research directions. Each paper is assigned one or multiple subject areas. The large number of subject areas makes it impossible to draw any conclusions in different research directions. Therefore, we use the classification in Wu et al. (2019) that merges WoS subject areas into 14 research fields, including physical sciences, chemistry, biology, medicine, agriculture, environmental and earth sciences, mathematics, computer and information technology, engineering, social sciences, business and management, law, humanities, and multidisciplinary sciences. This categorization is slightly different from what is recently proposed by Milojević (2020). However, because the observation is mainly in the field of natural science, the classification difference should not affect the results. A paper is usually tagged by multiple subject areas, it may also be labeled by multiple research fields. It is difficult to tell the priority in multiple subject areas; nor could we artificially tell which research field is closest to the content of the paper. Therefore, use whole counting to classify papers into research fields. In general, depending on the publication year, 20–25% of papers are labeled by multiple research fields.

2.7. Institution Diversity

WoS records the affiliation of each paper. Starting in 2008, it also records the affiliation of each author (i.e., who is affiliated with what institution). Therefore, there are two ways to analyze institution diversity. One is to use a paper's affiliations directly, and the other is to use the "main" affiliation of each author. Both approaches have pros and cons. Information about a paper's affiliation is easier to extract and is available for all papers in the data set. But given the trend that more authors are affiliated with multiple institutions (Hottenrott, Rose, & Lawson, 2019), directly using such information may overestimate institution diversity. In some cases we may also have a greater number of institutions than of authors. Using an author's "main" affiliation seems to be a more reasonable choice, which is also directly applied in the data set of Microsoft Academic Graph (Dong, Ma, et al., 2018; Wang, Shen, et al., 2020), but determining the primary affiliation out of others might be nontrivial. In this work, we use both methods to analyze institution diversity. If an author has multiple affiliations, we choose the one with the highest rank in the paper. We parse the institution information using the key value "organization" in the data, which usually refers to the university and the research lab. We report the results based on the author's affiliation in the main text. The results based on the paper's affiliation can be found in the Supplementary Information. The two approaches give slightly different results, but the conclusions drawn are the same. We are also aware of the name disambiguation issue in institution names (Donner, Rimmert, & van Eck, 2019). This should not affect our analyses because we compare institution in the same paper. It is very unlikely that authors would write the name of one institution in different ways in one single paper.

2.8. Funding Information

WoS started in 2008 to record the text related to funding acknowledgment for each paper, from where funding information, including the grant agency and grant ID, is parsed.

Despite concerns on the completeness and accuracy of the data (Álvarez-Bornstein, Morillo, & Bordons, 2017; Paul-Hus, Desrochers, & Costas, 2016; Tang, Hu, & Liu, 2017), it remains one of the largest available. We use such information directly as the criteria if a paper is funded. Because our measure is controlled by the national average value, we believe flaws in data recording should not affect the conclusion.

It is more complicated to search which paper is supported by the National Natural Science Foundation of China (NSFC) or the National Science Foundation (NSF) of the United States, because scientists acknowledge these funding agencies in different ways. For the NSFC, the most frequently used name is "National Natural Science Foundation of China," but other forms of the name, such as "Natural Science Foundation of China," "NSFC," "National Science Foundation of China," "National Nature Science Foundation of China," and "National Natural Science Foundation" are also widely used. The name variations of the NSF include "National Science Foundation," "NSF," and "National Science Foundation (NSF)." WoS has performed its own grant name disambiguation (which is available online), but such information is not available in our data set. Therefore, we extract the name of the grant agency in each paper from China and the United States, filter out those appearing fewer than 1,000 times in the data, and manually identify names associated with the NSFC and NSF. These names are list in Table S1 of the Supplementary Information. Other statistics given by our approach are listed in Table S2, which are in line with previous findings (Huang, Zhang, et al., 2016; Wang, Liu, et al., 2012).

It is noteworthy that the Ministry of Science and Technology (MOST) of China has its own research grants, such as the National Basic Research Program of China (973 Program), the National High Technology Research and Development Program of China (863 Program), and the National Key Technology R&D Program of China. The aim of these grants is to support big research groups (Figure S13). While they cover a relatively small fraction of scientific papers, the overlap with the NSFC is large. Around 17.5% of NSFC-supported papers are also supported by MOST, or equivalently 73.3% of MOST supported papers are simultaneously supported by the NSFC. To avoid potential bias, we remove papers that are supported by both NSFC and MOST and focus on those "primarily" supported by the NSFC. Note that the National Institutes of Health of the United States (NIH) also tends to support big groups (Figure S13). To make the comparison equal, we only consider papers "primarily" supported by the NSF by ignoring roughly 10.5% of NSF-supported papers that are also supported by the NIH. More statistics can be found in Table S3 of the Supplementary Information.

3. RESULTS

While collaboration plays an increasingly important role in scientific research, big teams have not yet taken over. In 2017, more than half of scientific papers were produced by teams with relatively small sizes (number of authors $m \leq 4$). The fraction of small team output differs from nation to nation, but China ranks last among the top 15 countries of scientific papers (Figures 1a and S1). In 2017, only 37% of papers from China are done by teams with $m \leq 4$, while this value is 58% for the United States and 55% for the global average (from which China is excluded).

We further analyze the $P(m \leq 4)$ in different research fields (Figures 1b, S1, S2, and S3). The statistics at the global level agree well with previous work and also with our intuition. For example, small teams are more frequently observed in mathematics, computer science, social science, business, humanities, and laws, with $P(m \leq 4)$ going beyond 80% or even higher. In interdisciplinary fields where collective intelligence is more important, $P(m \leq 4)$ drops to the
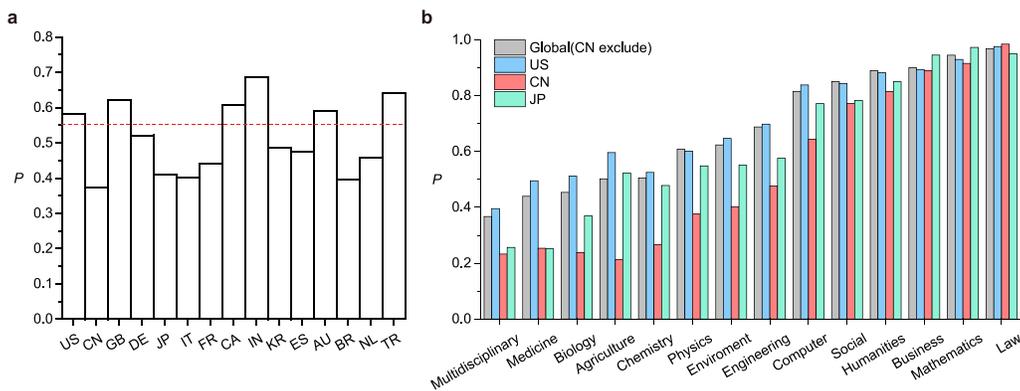
Figure 1. Panel a: The fraction of papers in 2017 produced by teams with size no more than 4 ($P(m \leq 4)$) in different countries. The dashed line corresponds to the global average (from which China is excluded). Panel b: $P(m \leq 4)$ in different fields in the year 2017.

lowest. Fields such as medicine, biology, chemistry, physics, and agriculture are usually believed to be labor intensive, requiring more individuals to be involved in research. But on the global average, $P(m \leq 4)$ is not very far below 50% and in some fields can even go above. Nevertheless, $P(m \leq 4)$ of China is significantly less than the global average in all areas of natural science. The relative difference is most prominent in agriculture, chemistry, biology, and medicine. In contrast, $P(m \leq 4)$ of the United States is greater than the global average in almost all areas of natural science. Being an Asian country with a large amount of scientific publications, Japan may be expected to be similar to China. But Japan's $P(m \leq 4)$ is closer to the global average and is much greater than that of China in all areas of natural science except medicine. In fields related to humanities, social science, and mathematics, $P(m \leq 4)$ of China is not very different from other countries (Li & Li, 2015), but papers in such fields make up only a very small fraction of China's annual production.

It is noteworthy that more papers being produced by big teams is a global trend. Indeed, we find in our analyses that the percentage of papers by small teams has decreased over years. Nevertheless, the drop for China is much steeper (Figures 2a and S4), giving rise to a statistically significant deviation from the global average (Supplementary Note 2). In 2000, $P(m \leq 4)$ of China is, though slightly smaller, not very different from that of the United States and the global average. But it goes down from 69.9% in 2000 to 37.4% in 2017, nearly 32 percentage points decrease. The drop, however, is only 17 percentage points for the United States (from 75.4% to 58.2%), 12 percentage points for Japan (from 53.0% to 41.0%), and 16 percentage points for the global average (from 70.8% to 55.2%). China's drop in small team output in fields of natural science and engineering is much higher than that of the global average (Figures 2b and S4), in line with our initial finding that small team output is small in these fields.

The observation that big teams dominate China's research output gives rise to another question: How would the team composition change when it shifts from small to large size? Indeed, a team can increase its size by adding more similar members or involving members with different backgrounds. While the team size grows in either way, team diversity is different, which has proved to be an essential factor in building a successful team (AlShebli, Rahwan, & Woon, 2018; Powell, 2018). There are different types of team diversity, such as ethnicity, discipline, gender, affiliation, and academic age (AlShebli et al., 2018; Huang, Gates, et al., 2020; Jia, Wang, & Szymanski, 2017). Here we focus on the affiliation and analyze the diversity at the
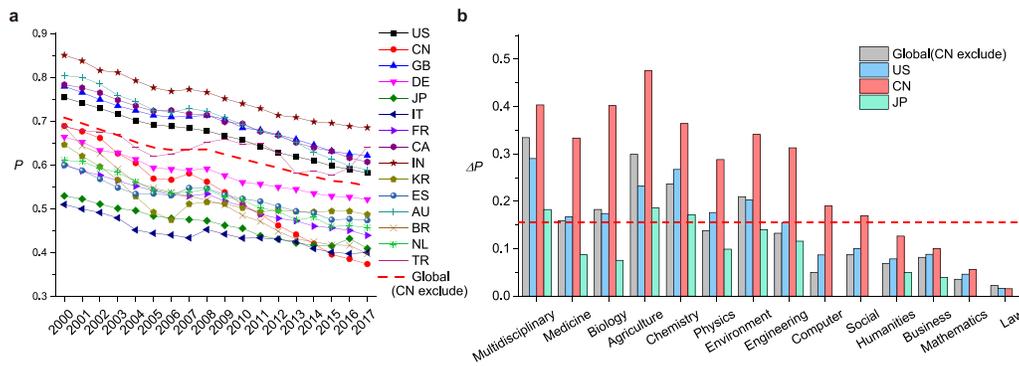
Figure 2. Panel a: The time evolution of $P(m \leq 4)$ in different countries. The dashed line corresponds to the global average (in which China is excluded). While it is a global trend that more papers are being produced by big teams, China's drop is much steeper. Panel b: The drop of $P(m \leq 4)$ from 2000 to 2017 in different fields. The dashed line corresponds to the drop of the global value in Panel a. China's drop is most prominent in fields of natural science and engineering, and can sometimes be twice as much as the global value. Note that the small team output has increased slightly in Japan in fields of social science and laws, giving rise to a negative value of ΔP. As we focus on the drop, we do not include this in the figure.

institution (organization) level. Indeed, a smaller team whose members are from diverse institutions is more likely to generate "hit" papers than a relatively larger team within one institution (Dong et al., 2018; Jones, Wuchty, & Uzzi, 2008). Here, we find that China's team composition is close to that of other countries when the team size is small, demonstrating a similar extent of institution diversity (Figures 3a and S5). However, different from other countries, China's institution diversity increases much more slowly as the team size increases. A significant fraction of big teams remain to be formed by members from the same institution (Figures 3b, 3c, and S5). For example, for all China's papers by six authors in 2017, nearly 50% of them are done in the same institution, which is 14 percentage points higher than for the United States. A similar conclusion also holds when we use the fraction of papers done by no more than two institutes. It is encouraging to notice the trend that teams tend to become more diverse as time goes by. One-institution papers take a smaller percentage of total publications now than in the past (Figure S6). However, the rate of change is low, suggesting that the institution diversity for China, an important factor for innovative work, will not improve very much in the near future.

So far, we have demonstrated the aspects in which China differs from other counties in research teams. What remain unclear are the factors that give rise to the differences observed. Given the confounding factors in teams assembling in different countries, identifying these factors is out of this paper's scope. Nevertheless, we perform some preliminary analyses by proposing and testing two hypotheses that seem capable of explaining the observations.

H1: Papers by big teams have a higher capability of receiving more citations in China than in other countries.
H2: Big teams are more preferred by funding agencies in China than in other countries.

Note that papers by larger teams on average receive more citations than those by smaller teams (Klug & Bagrow, 2016; Wu et al., 2019; Wuchty et al., 2007). The argument for H1 is that the citation boost is more considerable in China, which consequently provides incentives to build large teams. We count the total number of citations a paper receives within 5 years
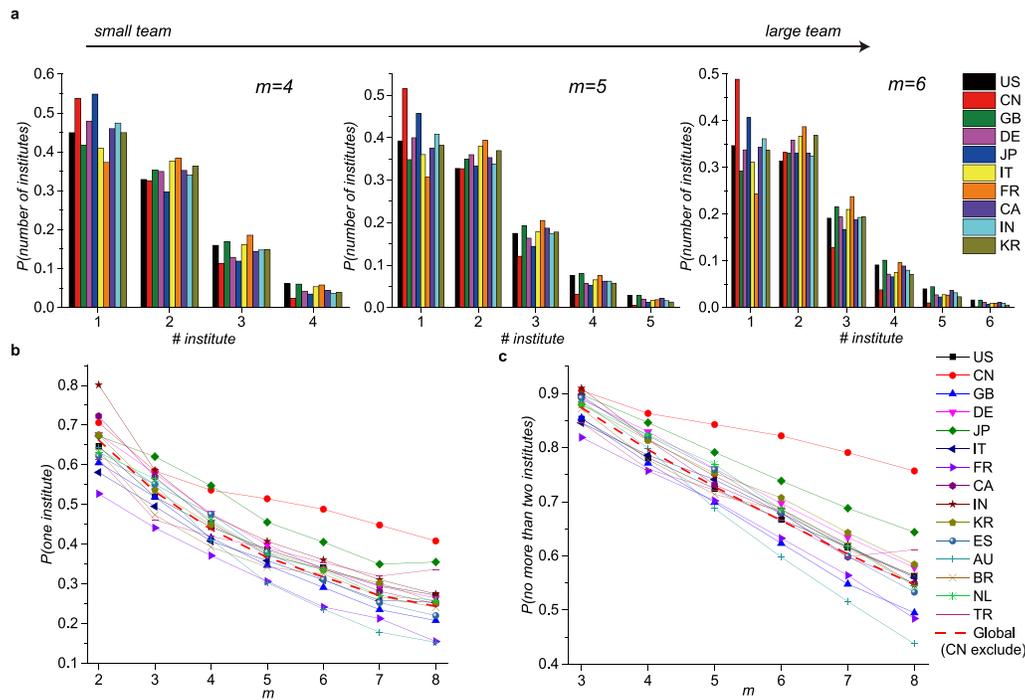
Figure 3. Panel a: The distribution of the number of distinct institutions in papers published in 2017 with team size $m = 4$, 5, and 6 (from left to right). China's team composition is not very different from other countries when the team is small. But as the team size increases, the distribution becomes more dominated by output from one institution. Panel b: The fraction of papers in 2017 done by one institution, given the team size m. A greater percentage of paper output is from a single institution in China than in other countries. Panel c: Similar to Panel b. The fraction of papers in 2017 involving no more than two institutions.

of its publication (c_5), and find that c_5 overall positively correlates with the team size m (Figure 4a). Papers from different countries receive different levels of citations. However, after re-scaling these curves by the national average of c_5, they almost collapse to a single curve (Figures 4b and S7). The trend that papers by bigger teams receive more citations is not different, or at least not more extreme, in China than in other countries. The same conclusion also holds when we use a shorter time window to count citations (Figure S8). Hence we conclude that H1 is not supported by the data.

We test H2 by extracting the grant information of each paper. Over 80% of papers from China contain grant information, much higher than for other countries (Figure S9). This implies that Chinese scientists are more obligated to acknowledge the funding agencies, or simply that only teams capable of securing research grants can efficiently conduct scientific research (Wang et al., 2012; Wang, Jones, & Wang, 2019; Yang, Gu, Wang, Hu, & Tang, 2015). Either of these explanations suggests the significant impact of funding agencies on scientific research in China. As intuitively expected, the percentage of papers with grants increases with team size in almost every country (Figure S9). But once again, the increase is not sharper in China than in other countries (Figure S9). Nor could we find any difference in the number of grants a paper is supported by (Figure S10).

Indeed, given different sources of funding in different nations, different policies and aims of different funding agencies, and potential flaws in the records that may affect the observation (Álvarez-Bornstein et al., 2017; Azoulay, Graff Zivin, & Manso, 2011; Paul-Hus et al., 2016; Tang et al., 2017), it would be less meaningful to test H2 by comparing all grants and papers
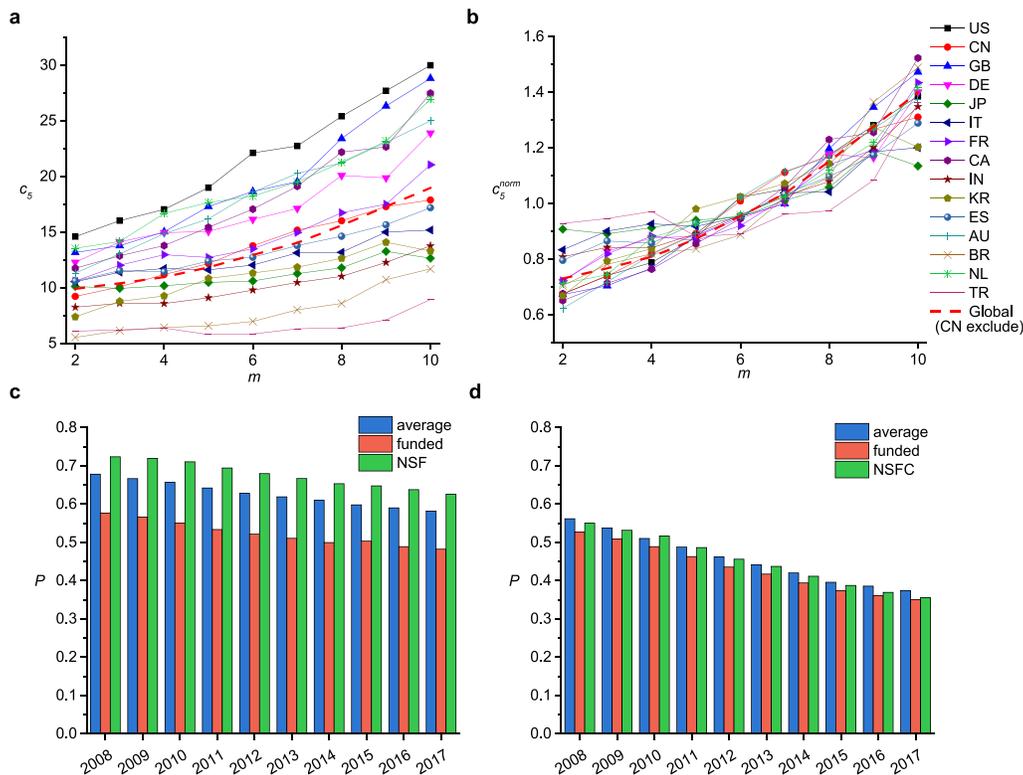
Figure 4. Panel a: The total number of citations a paper receives within five years of its publication c_5 is positively correlated with the team size m in every country. The statistics are based on papers published in 2011. Panel b: When re-scaling the number of citation of c_5 by the average value of a country as $c_5^{norm} = c_5/\langle c_5 \rangle$, different curves in Panel a almost collapse to a single curve showing similar increasing trend with team size. Panel c: The fraction of small team output ($m \leq 4$) among all papers, funded papers, and papers supported by the NSF from the United States. While funding agencies in general prefer big teams, the NSF supports more small team work than the average. Panel d: The fraction of small team output ($m \leq 4$) among all papers, funded papers, and papers supported by the NSFC from China. In 2011, the NSFC supports slightly more small team works than the national average. In 2010, the difference between the NSFC and the national average is not statistically significant. Most of the time, the NSFC supports less small team work than the average.

from all countries. For this reason, we then consider only two granting agencies: the National Natural Science Foundation of China (NSFC) and the National Science Foundation of the United States (NSF). It is believed that China learned from the NSF in initiating and organizing the NSFC. The two have very similar budgets (especially after taking purchasing power into consideration), scope, and aim. In addition, both of them are one of the major national funding resources for fundamental research (Huang et al., 2016; Wang et al., 2012). All these features make the NSFC and NSF two comparable examples. For each of China and the United States, we collect three sets of papers: all papers published in a given year, papers supported by grants in that year, and papers mainly supported by the NSFC or NSF in that year (see Section 2 for details). Compared with the national average, small team output is lower in papers with grants (Figures 4c, 4d, S11, and S12), in line with our previous finding that works by larger teams have a higher probability of being sponsored. However, within papers supported by the NSF, the percentage of output from small teams is higher than average (Figures 4c, S11, and S12). In contrast, the fraction of small team output is usually less than average in papers supported by the NSFC (Figures 4d, S11, and S12). In other words, using the national average as the baseline, the NSF supports more small team work than the NSFC does. Given the similarities between the two funding agencies, this observation supports H2.

It may be argued that the NSFC and NSF are not comparable because China does not have an independent funding agency like the NIH that mainly focuses on biomedical research. Therefore, the NSFC supports more work in medicine that relies mainly on cooperation by big teams. Consequently, the percentage of small team output is dragged down. Statistically, the argument stands. The fraction of supported works in biology is roughly the same for the NSFC and NSF, where 12% of NSFC-supported work and 13.5% of NSF-supported work is in biology. But there is a nonnegligible difference in the field of medicine: 9.6% of NSFC-supported work is in medicine, while this value is only 2.4% for the NSF. Such a difference by itself is related to intriguing questions in research management and policy, as it is unclear if combining application-oriented research such as medicine with basic research, like the NSFC does, would enhance the efficiency. Nevertheless, in terms of data analyses, we can treat the data by excluding NSFC and NSF-supported papers in the field of medicine. After this modification, papers supported by the NSFC that are carried out by small teams are slightly more than the national average (Figures S14 and S15). The extent to which the NSFC is over the national average, however, is still smaller than that of the NSF. Hence, even after excluding papers in medicine, the NSF supports more small team work than the NSFC does, supporting our conclusion above.

4. CONCLUSION

To summarize, we analyze over 26 million papers on WoS published from 2000 to 2017, which is one of the most extensive analyses in terms of papers covered. We find that China's research output is more dominated by big teams than the rest of the world. The fraction of papers by small teams in China is not only much lower than the global average, ranking last among the top 15 countries of scientific publications in 2017, but has also undergone a much steeper decrease since 2000. More importantly, as teams in China shift from small to large size, the team diversity that is essential for innovative work does not follow the same increase as that in other countries. A high percentage of work is carried out within one or two institutions. All of these observations indicate that China is very different from other countries in the composition of big and small teams in scientific research. If referring to the global average or a country like the United States, China is a long way from the balance point.

Given the importance of the problem, we also make some preliminary attempts to understand factors that explain the different small/big team composition in China. The first hypothesis we test is that China's big teams have a more considerable advantage in gaining citations than those of other countries. Hence, there are more incentives to build a large team. Indeed, work by larger teams on average receives more citations than those by small teams. However, the citation boost is roughly the same in every country after taking the national average citation into consideration, implying that citation alone cannot explain the difference. We then turn to checking whether large teams are more preferred by funding agencies. More than 80% of papers from China acknowledge research grants, which is the highest among the 15 countries analyzed. It clearly indicates the significant influence of funding agencies on China's scientific research. While works by large teams are more apt to be supported by grants, China does not demonstrate any different patterns in this matter, following the same trend as in other countries. Yet, when we separately compare the work supported by the NSFC and NSF, we find that the NSFC supports less small-team work than the NSF does. This gives some clues supporting the hypothesis that preferences by the funding agencies may be associated with the imbalance of small and big teams in China.

Concern about the balance between small and big teams is a relatively new topic that has rarely been studied in the past. Nevertheless, if we admit that small and large teams play different yet equally essential roles in scientific research, we need to consider the imbalance

seriously. Our analyses, based on the large volume of publication data, provides evidence suggesting that China may need more small team output. If the fall in small teams persists, China may become less competitive in delivering disruptive research outcomes and expanding the frontier of the field. One day, the science community in China may not have enough new questions for its big teams to further develop and exploit. The factor we have spotted that is associated with this imbalance further sheds light on this issue. Giving multiple confounding factors that may influence the organization of teams, we admit that our finding is preliminary. For example, one fundamental assumption of this study is that the patterns observed in big and small teams are universal. It is, however, reasonable to question this assumption. Indeed, if China's big teams are as capable as small teams in performing disruptive research, or if team diversity is not correlated with the impact of the work in China, the results reported in this paper would raise little worry. Currently we have some preliminary results confirming that patterns in big and small teams are universal, providing the basis of the research. But nationality and universality in the science of science study is an interesting future direction. Some observations in this paper can be explained by the fact that big teams in China are more productive than those in other countries. However, given the fluid nature of team assembly (Abramo, D'Angelo, & Murgia, 2017; Milojević, 2014; Wang & Hicks, 2015), testing this hypothesis is challenging, which requires a better author name disambiguation algorithm and other techniques to extract the core of the team in the scientific collaboration network (Wang, Ran, & Jia, 2020; Yu, Xia, & Liu, 2019). The lack of team diversity and the large average team size in China can also be associated with the honorary authorship, where scholars not directly contributing to the work are added to the author list (Biagioli, Kenney, et al., 2019). Although there is no evidence that such misconduct is more severe in China (Tang, 2019), the effects of honorary authorship on the team size may be worth further investigations. The collectivist culture in Asia may also encourage the formation of big teams. Both Japan and Korea have a relatively small percentage of small team output. Exploring these factors may not only provide useful insights to the research community in China but also advance our quantitative understanding of science (Azoulay et al., 2018; Fortunato et al., 2018).

AUTHOR CONTRIBUTIONS

Linlin Liu: Methodology, Software, Formal analysis, Investigation, Visualization, Writing—Original draft. Jianfei Yu: Software, Investigation, Validation. Junming Huang: Conceptualization, Data Curation, Writing—Review & editing. Feng Xia: Conceptualization, Writing—Review & editing. Tao Jia: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing—Original Draft, Writing—Review & editing, Supervision.

COMPETING INTERESTS

The authors have no competing interests.

DATA AVAILABILITY

The data used in this paper are proprietary and cannot be posted in a repository.

REFERENCES

Abramo, G., D'Angelo, A. C., & Murgia, G. (2017). The relationship among research productivity, research collaboration, and their determinants. *Journal of Informetrics*, 11(4), 1016–1030. **DOI:** https://doi.org/10.1016/j.joi.2017.09.007

AlShebli, B. K., Rahwan, T., & Woon, W. L. (2018). The preeminence of ethnic diversity in scientific collaboration. *Nature Communications*, 9(1), 1–10. **DOI:** https://doi.org/10.1038/s41467-018-07634-8, **PMID:** 30514841, **PMCID:** PMC6279741

Álvarez-Bornstein, B., Morillo, F., & Bordons, M. (2017). Funding acknowledgments in the Web of Science: Completeness and accuracy of collected data. *Scientometrics*, 112(3), 1793–1812. **DOI:** https://doi.org/10.1007/s11192-017-2453-4

Azoulay, P. (2019). Small-team science is beautiful. *Nature*, 566(7744), 330–332. **DOI:** https://doi.org/10.1038/d41586-019-00350-3, **PMID:** 30783269

Azoulay, P., Graff-Zivin, J., Uzzi, B., Wang, D., Williams, H., … Guinan, E. C. (2018). Toward a more scientific science. *Science*, 361(6408), 1194–1197. **DOI:** https://doi.org/10.1126/science.aav2484, **PMID:** 30237341

Azoulay, P., Graff Zivin, J. S., & Manso, G. (2011). Incentives and creativity: evidence from the academic life sciences. *RAND Journal of Economics*, 42(3), 527–554. **DOI:** https://doi.org/10.1111/j.1756-2171.2011.00140.x

Barlow, J., Stephens, P. A., Bode, M., Cadotte, M. W., Lucas, K., … Pettorelli, N. (2018). On the extinction of the single-authored paper: the causes and consequences of increasingly collaborative applied ecological research. *Journal of Applied Ecology*, 55(1), 1–4. **DOI:** https://doi.org/10.1111/1365-2664.13040

Biagioli, M., Kenney, M., Martin, B., & Walsh, J. P. (2019). Academic misconduct, misrepresentation and gaming: A reassessment. *Research Policy*, 48(2), 401–413. **DOI:** https://doi.org/10.1016/j.respol.2018.10.025

Castelvecchi, D. (2015). Physics paper sets record with more than 5,000 authors. *Nature*, 15. **DOI:** https://doi.org/10.1038/nature.2015.17567

Cummings, J. N., & Kiesler, S. (2007). Coordination costs and project outcomes in multi-university collaborations. *Research Policy*, 36(10), 1620–1634. **DOI:** https://doi.org/10.1016/j.respol.2007.09.001

Dong, Y., Ma, H., Tang, J., & Wang, K. (2018). Collaboration diversity and scientific impact. arXiv preprint *arXiv:1806.03694*.

Donner, P., Rimmert, C., & van Eck, N. J. (2019). Comparing institutional-level bibliometric research performance indicator values based on different affiliation disambiguation systems. *Quantitative Science Studies*, 1(1), 150–170. **DOI:** https://doi.org/10.1162/qss_a_00013

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., … Barabási, A.-L. (2018). Science of science. *Science*, 359(6379), eaao0185. **DOI:** https://doi.org/10.1126/science.aao0185, **PMID:** 29496846, **PMCID:** PMC5949209

Gazni, A., Sugimoto, C. R., & Didegah, F. (2012). Mapping world scientific collaboration: Authors, institutions, and countries. *Journal of the American Society for Information Science and Technology*, 63(2), 323–335. **DOI:** https://doi.org/10.1002/asi.21688

González-Alcaide, G., Park, J., Huamaní, C., & Ramos, J. M. (2017). Dominance and leadership in research activities: Collaboration between countries of differing human development is reflected through authorship order and designation as corresponding authors in scientific publications. *PLoS ONE*, 12(8), e0182513. **DOI:** https://doi.org/10.1371/journal.pone.0182513, **PMID:** 28792519, **PMCID:** PMC5549749

Guo, J., Liu, X., Yang, L., & Wu, J. (2019). Are contributions from Chinese physicists undercited? *Journal of Data and Information Science*, 4(4), 84–95. **DOI:** https://doi.org/10.2478/jdis-2019-0022

Hottenrott, H., Rose, M., & Lawson, C. (2019). The rise of multiple institutional affiliations. arXiv preprint *arXiv:1912.05576*. **DOI:** https://doi.org/10.2139/ssrn.3697216

Huang, F. (2018). Quality deficit belies the hype. *Nature*, 564(7735), S70–S71. **DOI:** https://doi.org/10.1038/d41586-018-07694-2, **PMID:** 30542188

Huang, J., Gates, A. J., Sinatra, R., & Barabasi, A.-L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences*, 117(9), 4609–4616. **DOI:** https://doi.org/10.1073/pnas.1914221117, **PMID:** 32071248, **PMCID:** PMC7060730

Huang, M.-H., Lin, C.-S., & Chen, D.-Z. (2011). Counting methods, country rank changes, and counting inflation in the assessment of national research productivity and impact. *Journal of the American Society for Information Science and Technology*, 62(12), 2427–2436. **DOI:** https://doi.org/10.1002/asi.21625

Huang, Y., Zhang, Y., Youtie, J., Porter, A. L., & Wang, X. (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. **DOI:** https://doi.org/10.1371/journal.pone.0154509, **PMID:** 27219466, **PMCID:** PMC4878788

Jia, T., Wang, D., & Szymanski, B. K. (2017). Quantifying patterns of research-interest evolution. *Nature Human Behaviour*, 1(4), 1–7. **DOI:** https://doi.org/10.1038/s41562-017-0078

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *Review of Economic Studies*, 76(1), 283–317. **DOI:** https://doi.org/10.1111/j.1467-937X.2008.00531.x

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-university research teams: Shifting impact, geography, and stratification in science. *Science*, 322(5905), 1259–1262. **DOI:** https://doi.org/10.1126/science.1158357, **PMID:** 18845711

Kahn, S., & MacGarvie, M. (2016). Do return requirements increase international knowledge diffusion? evidence from the Fulbright program. *Research Policy*, 45(6), 1304–1322. **DOI:** https://doi.org/10.1016/j.respol.2016.02.002

Kao, C. (2009). The authorship and country spread of operation research journals. *Scientometrics*, 78(3), 397–407. **DOI:** https://doi.org/10.1007/s11192-008-1850-0

Klug, M., & Bagrow, J. P. (2016). Understanding the group dynamics and success of teams. *Royal Society Open Science*, 3(4), 160007. **DOI:** https://doi.org/10.1098/rsos.160007, **PMID:** 27152217, **PMCID:** PMC4852640

Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*, 66(7), 1323–1332. **DOI:** https://doi.org/10.1002/asi.23266

Larsen, P. (2008). The state of the art in publication counting. *Scientometrics*, 77(2), 235–251. **DOI:** https://doi.org/10.1007/s11192-007-1991-6

Leahey, E. (2016). From sole investigator to team scientist: Trends in the practice and study of research collaboration. *Annual Review of Sociology*, 42, 81–100. **DOI:** https://doi.org/10.1146/annurev-soc-081715-074219

Lewison, G., Purushotham, A., Mason, M., McVie, G., & Sullivan, R. (2010). Understanding the impact of public policy on cancer research: A bibliometric approach. *European Journal of Cancer*, 46(5), 912–919. **DOI:** https://doi.org/10.1016/j.ejca.2009.12.020, **PMID:** 20064708

Li, J., & Li, Y. (2015). Patterns and evolution of coauthorship in China's humanities and social sciences. *Scientometrics*, 102(3), 1997–2010. **DOI:** https://doi.org/10.1007/s11192-014-1471-8

Lin, C.-S., Huang, M.-H., & Chen, D.-Z. (2013). The influences of counting methods on university rankings based on paper count and citation count. *Journal of Informetrics*, 7(3), 611–621. **DOI:** https://doi.org/10.1016/j.joi.2013.03.007

Mazloumian, A., Helbing, D., Lozano, S., Light, R. P., & Börner, K. (2013). Global multi-level analysis of the 'scientific food web'. *Scientific Reports*, 3(1), 1–5. **DOI:** https://doi.org/10.1038/srep01167, **PMID:** 23378902, **PMCID:** PMC3558694

Milojević, S. (2014). Principles of scientific research team formation and evolution. *Proceedings of the National Academy of Sciences*, 111(11), 3984–3989. **DOI:** https://doi.org/10.1073/pnas.1309723111, **PMID:** 24591626, **PMCID:** PMC3964124

Milojević, S. (2020). Practical method to reclassify Web of Science articles into unique subject categories and broad disciplines. *Quantitative Science Studies*, 1(1), 183–206. **DOI:** https://doi.org/10.1162/qss_a_00014

Newman, M. E. (2001). Scientific collaboration networks. I. network construction and fundamental results. *Physical Review E*, 64(1), 016131. **DOI:** https://doi.org/10.1103/PhysRevE.64.016131, **PMID:** 11461355

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. **DOI:** https://doi.org/10.1007/s11192-016-1953-y

Phillips, N. (2016). China: Building an innovator. *Nature*, 533(7601), S32–S33. **DOI:** https://doi.org/10.1038/533S32a, **PMID:** 27144607

Powell, K. (2018). These labs are remarkably diverse—here's why they're winning at science. *Nature*, 558, 19–22. **DOI:** https://doi.org/10.1038/d41586-018-05316-5, **PMID:** 29875493

Price, D. J. (1963). *Little science, big science* (Vol. 5). New York: Columbia University Press. DOI: https://doi.org/10.7312/pric91844

Sivertsen, G., Rousseau, R., & Zhang, L. (2019). Measuring scientific contributions with modified fractional counting. *Journal of Informetrics*, 13(2), 679–694. **DOI:** https://doi.org/10.1016/j.joi.2019.03.010

Tang, L. (2019). Five ways China must cultivate research integrity. *Nature*, 575, 589–591. **DOI:** https://doi.org/10.1038/d41586-019-03613-1, **PMID:** 31768041

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. **DOI:** https://doi.org/10.1002/asi.23713

Thelwall, M. (2019). Large publishing consortia produce higher citation impact research but coauthor contributions are hard to evaluate. *Quantitative Science Studies*, 1(1), 290–302. **DOI:** https://doi.org/10.1162/qss_a_00003

Waltman, L., & van Eck, N. J. (2015). Field-normalized citation impact indicators and the choice of an appropriate counting method. *Journal of Informetrics*, 9(4), 872–894. **DOI:** https://doi.org/10.1016/j.joi.2015.08.001

Wang, J., & Hicks, D. (2015). Scientific teams: Self-assembly, fluidness, and interdependence. *Journal of Informetrics*, 9(1), 197–207. **DOI:** https://doi.org/10.1016/j.joi.2014.12.006

Wang, K., Shen, Z., Huang, C.-Y., Wu, C.-H., Dong, Y., & Kanakia, A. (2020). Microsoft academic graph: When experts are not enough. *Quantitative Science Studies*, 1(1), 396–413. **DOI:** https://doi.org/10.1162/qss_a_00021

Wang, X., Liu, D., Ding, K., & Wang, X. (2012). Science funding and research output: a study on 10 countries. *Scientometrics*, 91(2), 591–599. **DOI:** https://doi.org/10.1007/s11192-011-0576-6

Wang, X., Ran, Y., & Jia, T. (2020). Measuring similarity in co-occurrence data using ego-networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 30(1), 013101. **DOI:** https://doi.org/10.1063/1.5129036, **PMID:** 32013468

Wang, Y., Jones, B. F., & Wang, D. (2019). Early-career setback and future career impact. *Nature Communications*, 10(1), 1–10. **DOI:** https://doi.org/10.1038/s41467-019-12189-3, **PMID:** 31575871, **PMCID:** PMC6773762

Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, 566(7744), 378–382. **DOI:** https://doi.org/10.1038/s41586-019-0941-9, **PMID:** 30760923

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316(5827), 1036–1039. **DOI:** https://doi.org/10.1126/science.1136099, **PMID:** 17431139

Xie, Y., Zhang, C., & Lai, Q. (2014). China's rise as a major contributor to science and technology. *Proceedings of the National Academy of Sciences*, 111(26), 9437–9442. **DOI:** https://doi.org/10.1073/pnas.1407709111, **PMID:** 24979796, **PMCID:** PMC4084436

Yang, X., Gu, X., Wang, Y., Hu, G., & Tang, L. (2015). The Matthew Effect in China's science: Evidence from academicians of Chinese Academy of Sciences. *Scientometrics*, 102(3), 2089–2105. **DOI:** https://doi.org/10.1007/s11192-014-1502-5

Yu, S., Xia, F., & Liu, H. (2019). Academic team formulation based on Liebig's barrel: Discovery of anticask effect. *IEEE Transactions on Computational Social Systems*, 6(5), 1083–1094. **DOI:** https://doi.org/10.1109/TCSS.2019.2913460

Zheng, J., Zhao, Z., Zhang, X., Huang, M.-H., & Chen, D.-Z. (2014). Influences of counting methods on country rankings: A perspective from patent analysis. *Scientometrics*, 98(3), 2087–2102. **DOI:** https://doi.org/10.1007/s11192-013-1139-9

Zhou, Y., Lazonick, W., & Sun, Y. (2016). *China as an innovation nation*. Oxford: Oxford University Press. **DOI:** https://doi.org/10.1093/acprof:oso/9780198753568.001.0001