



# What is the relationship between research funding and citation-based performance? A comparative analysis between critical disciplines

Saeed Roshani<sup>1</sup> · Mohammad-Reza Bagheryllooieh<sup>1</sup> · Melika Mosleh<sup>2</sup> · Mario Coccia<sup>3</sup>

Received: 16 December 2020 / Accepted: 9 June 2021 / Published online: 24 June 2021  
© Akadémiai Kiadó, Budapest, Hungary 2021

## Abstract

The goal of the study here is to model and analyze the relation between research funding and citation-based performance in science to predict the diffusion of new scientific results in society. In fact, an important problem in the field of scientometrics is to explain factors determining the growth of citations in documents that can increase the diffusion of scientific results and the impact of science on society. The study here confronts this problem by developing a scientometric analysis to clarify, whenever possible, the relation between research funding and citations of articles in critical disciplines. Data of 2015 retrieved from the Web of Science database relating to the three critical disciplines given by computer science, medicine and economics are analyzed. Results suggest that computer science journals published more funded than unfunded papers. Medicine journals published equally funded and unfunded documents, and finally economics journals published more unfunded than funded papers. In addition, funded documents received more citations than unfunded papers in all three disciplines under study. The study also finds that citations in funded, unfunded and total (funded + unfunded) papers follow a power-law distribution in different disciplines. Another novel finding is that for all disciplines under study, the Matthew effect is greater for funded articles compared to unfunded documents. The results here can support best practices of research policy directed to fund vital scientific research for increasing the diffusion of science and scientific findings in society.

**Keywords** Research funding · Science creation · Science diffusion · Citations · Scientific development · Papers · Power-law distribution · Matthew effect

## Introduction

Many studies have attempted to explore the impacts of different variables on research outputs (Rousseau, 2000; Van Raan, 1998). Some researches focus on the relationship between scientific collaboration and research impact (cf., Coccia & Bozeman, 2016; Ronda-Pupo

---

✉ Saeed Roshani  
Roshani@atu.ac.ir

Extended author information available on the last page of the article

& Katz, 2016a, 2016b). Other papers explore the scaling relationship between the size of the innovation system and citation-based performance (Katz, 2016; Ronda-Pupo, 2017). In this context, funding plays a critical role in the quality and performance of scientific studies, such as on the impact factor of journals (Wang & Shapira, 2015; Lyall et al. 2013). Although some studies do not find any relationship between funding and first citation, there can be a strong link between funding and top percentile citations across different funding sources (Gök et al. 2016). Huang et al. (2006) argue that a significant relationship exists between patent citations and NSF-funded researchers. Wang and Shapira (2015) identify a relationship between research impact and research sponsorship in nanotechnology papers. In particular, they suggest that sponsored research publications have a higher impact than researches without grant sponsors. Quinlan et al. (2008) confirm funded publications receive more citations compared to unfunded research. In computer science, Shen et al. (2016) show a connection between funding and the citations of research. Yan et al. (2018) also show that funded publications in science, technology, engineering, and mathematics (STEM), multi-author, and multi-institution papers tend to receive more citations than single-authored and single-institution researches. Other studies do not show significant differences in terms of scientific impact between funded and unfunded research (Jacob and Lefgren, 2011; Rigby, 2013). Cronin and Shaw, (1999) maintain that the citation of articles is associated with journals and the author's nationality in four information science journals. Hicks and Katz, (2011) argue that in science, research funding should be based on the merit-based evaluation of scholars, but there is unequal distribution of performance in science, and the decision-makers in power face a lot of difficulties in supplying public funds to overcome the level of inequality in research performance. Zhao et al. (2018) argue that research funding is one of the most important resources in the reward system of science. They suggest that funding has a main impact on usage and citations, and funded papers can attract more usage in different fields of research (Morillo, 2020; Pao, 1991). In general, the literature suggests that funded studies are likely to enjoy a higher research impact than other publications.

However, in scientometrics the investigation of the scaling relationship between funding of research and impact of paper in terms of higher citations between different disciplines is hardly known. Stimulated by these fundamental problems, the main goal of the present study is to verify the presence of power-law distribution between the citation-based performance of articles and their funding status in three main disciplines (computer science, medicine and economics) in order to suggest empirical properties that can explain and generalize whenever possible factors associated with the impact of science in society. These topics of 'the science of science' can offer a deeper understanding of successful driving factors of science to support best practices of research policy (Coccia, 2018, 2020; Coccia & Wang, 2016; Fortunato et al. 2018).

## Theoretical framework

In order to position our analysis in scientific literature, we begin by reviewing existing approaches and studies. A power-law approach implies an inverse exponential relationship between two variables (cf. Barabási & Albert, 1999). Scaling behavior in bibliometrics is based on the Lotka, (1926) rule, which describes the '20/80' distribution of scientific productivity through authors and articles. Cumulative advantage (de Solla-Price, 1976) and success-breeds-success (Merton, 1968; Merton, 1988) are two main concepts associated

with many power-laws and scaling behavior. Lotka, (1926) claimed that only a few authors account for around 80 percent of scientific findings. Scholars have attempted to provide evidence that the distributions of other aspects in science (e.g. citations, number of articles in journals, word counts, number of authors for each paper, etc.) can have a power-law distribution (Ye & Rousseau, 2008). For instance, de Solla-Price (1976) argues that the probability of receiving citations for a paper depends on previous citations, such that scientific papers with the highest number of citations tend to be cited more quickly than papers with a small number of citations. In particular, the number of citations as a function of the number of publications tends to produce a heavy-tailed distribution illustrated by power-law. Katz (1999, 2000, 2005) studied scaling relationships between the number of publications and citations across research institutes, various countries, and disciplines (cf. Coccia, 2005a, b, 2019). Katz (2000, 2005) provides strong evidence that the scientific system is characterized by a size-dependent cumulative advantage: a non-linear increase of impact is led by a growth in size. It is assumed that most publications receive the least citations while an only a few papers enjoy a significant impact, becoming the main node of the citation network (cf. van Raan, 2006, 2008). Ronda-Pupo and Katz, (2018) have evaluated the power-law correlation between multi-authorship and citation impact in articles of information science & library science journals, revealing the existence of the Matthew effect among some of the scientific resource elements. Wang and Shapira, (2011) and Gao, (2019) have analyzed the distribution of citations of over-funded publications. Scholars also suggest two models that indicate the scale-invariance property (Katz, 2005; Ronda-Pupo, 2017). The first one is a power-law probability distribution, and the second is the power-law correlations. A scaling relationship exists between two entities ( $x$  and  $y$ ) if they are correlated by a power-law function (Ronda-Pupo & Katz, 2018).

The present paper empirically analyses the power-law distribution of citations of total (funded + unfunded)/funded/unfunded studies and the power-law correlation between the number of total (funded + unfunded)/funded/unfunded papers and their citations impact in different disciplines. The scaling exponent approach will be used in the present study to explore whether any scale-invariant measures and models can aid the decision-making of policymakers to invest in different disciplines in order to predict and increase the diffusion of scientific results and their impact on society.

## Materials and methods

### Research questions and motivation of the study

The main research questions of this study are:

- How does research funding affect citations of papers in different disciplines?
- Do the number of citations in funded, unfunded and total (funded + unfunded) papers follow a power-law distribution?
- Is the growth of the citations in funded research different from unfunded research?

We confront these questions here by developing an inductive study, which endeavors to explain, whenever possible, common patterns of the distribution of citations of funded and unfunded articles in different disciplines, and how research funding affects the citations and as a consequence the diffusion of scientific research. The development of the study

is based on a comparative analysis of three disciplines, given by computer science, medicine, and economics, to analyze similarities and/or differences in the relationship between research funding and citations and to clarify article-level success and failure. We are interested in providing best practices of research funding for universities and at a national level to foster the diffusion of good research (Coccia, 2005a, 2019).

### Data, sources and retrieval strategy

- Disciplines under study here are computer science, medicine and economics. These disciplines fields are representative of critical research fields in different sciences (cf. Coccia & Wang, 2016).
- The data of the study are all publications of the year 2015 in the fields of computer science, medicine, and economics from the Web of Science database, including inclusive and citation counts for six fixed citation windows (Web of Science, 2021).

We used the publication types of articles and reviews that were published in journals. The string search for retrieving computer science data was WC= "Computer Science", refined by document types: (Articles OR Review). We changed the WC value to "Medicine" and "Economics" to retrieve data related to the fields of medicine and economics. Although a significant number of publications in the field of computer science were conference proceedings, the study did not consider these documents as less than ten percent of them were funded, creating distortion in the comparative analysis done here. For the field of medicine, the number of funded and unfunded papers were roughly equal, providing a representative power. Funding research in social sciences were scarce compared to the two other fields under study here (Coccia et al. 2015). Accordingly, we chose the field of economics in which the portion of funded papers was significant enough to be analyzed in this study.

### Measures

The number of papers in computer science, medicine and economics was the explanatory variable taken from the Web of Science database (2021).

Citation-based performance was measured with the number of citations in three different sets: funded articles, unfunded articles and the total (funded and unfunded) articles between disciplines under study (Web of Science, 2021).

This study focused on the papers published in 2015 to consider a time lag of more than five years for assessing a comprehensive amount of citations to support theoretical and empirical analysis.

### The theoretical and empirical strategy of the model and the data analysis procedure

The analysis of research questions is based on three main steps.

Firstly, we analyzed the power-law distribution of the citations of total (funded + unfunded)/funded/unfunded articles published from computer science, medicine, and economics.

Secondly, we compared the power-law distribution fitness with other plausible distributions.

Thirdly, we analyzed the power-law correlation between citations and numbers of total (funded + unfunded)/funded/unfunded papers.

### Verification of the power-law distribution

In order to test the existence of the power-law distribution of the citations of total (funded + unfunded)/funded/unfunded articles in all documents, we followed the procedure developed by Clauset et al. (2009, p.3, box. 1). At first, we estimated the  $X_{\min}$  value and scaling parameter  $\alpha$  of the power-law. According to Clauset et al. (2009),  $X_{\min}$  is the value when the power law begins. Given that the data are discrete, to calculate the  $X_{\min}$  value, we used the formula devised by Clauset et al. (2009):

$$\hat{\alpha} \simeq 1 + n \left[ \sum_{i=1}^n \ln \frac{x_i}{x_{\min} - \frac{1}{2}} \right]^{-1}$$

To calculate the parameter  $\alpha$ , we used the maximum likelihood method. After that, we calculated the goodness of fit between the data and power-law. The aim here is to verify whether the power-law distribution can be a suitable distribution of data for the disciplines under study here. In this step, we used the bootstrap function and ran 2500 Monte-Carlo simulations. We fitted each sample individually to its own power-law model and calculated the Kolmogorov–Smirnov (KS) statistics for each item in relation to its model. The Kolmogorov–Smirnov (KS) statistic quantifies a distance between the sample’s empirical distribution function and the reference’s cumulative distribution function.

The aim of this step was to determine the  $p$ -value. According to Clauset et al. (2009), if the  $p$ -value is greater than 0.1 in power-law distribution, the power-law is a suitable distribution for the scatter of data. In addition, we compared the power-law with alternative distributions. The primary purpose was to verify that the power-law distribution was the best fitting model compared to alternative distributions. In this step, we compared the power-law distribution to lognormal, exponential, Poisson, stretched exponential, and power-law with exponential cut-off distributions as alternative options. We also used the Kolmogorov–Smirnov (KS) test to measure the distance between the distributions. Then, we calculated the  $p$ -value for each of the distributions and compared them with the  $p$ -value of the power-law distribution. According to Clauset et al. (2009), the likelihood ratio (LR) is a determinant factor to choose between two alternative distributions. The distribution with the most negative LR and a significant  $p$ -value is a better fit than the power-law distribution. If all the alternative distributions have a positive LR, the power-law distribution can be considered as a proper fit for the model. All of these calculations were performed by Power-Law packages in the R software that was proposed by Gillespie, (2015). We used the Python package as proposed by Alstott et al. (2014) to analyze the stretched exponential and power-law with cut-off distributions. As Clauset et al. (2009) offered, we used key terms to assess the statistical analyses for the power-law distribution of each data set:

- ‘None’ indicates that the data is probably not power-law distributed. It means that the  $p$ -value of the power-law distribution is not significant.

- ‘Moderate’ indicates that the power-law is a good fit but that there are other plausible alternatives as well. It means that the power-law and other distributions have significant  $p$ -value, but the LR of the alternatives is positive.
- ‘Good’ indicates that the power-law is a good fit and that none of the alternatives considered are plausible. It means that just the power-law  $p$ -value is significant among all of the other distributions.
- ‘Cut-off’ indicates that the power-law with exponential cut-off is favored over the pure power-law. It means that the  $p$ -value of the power-law with exponential cut off is significant and has the most negative LR.

### Scaling correlation

Finally, we analyzed the power-law correlation between the number of citations and the number of total (funded + unfunded)/funded/unfunded papers in all articles published in computer science, medicine and economics. Citation-based performance (CBP) refers to all citations received by the articles within computer science, medicine and Economics. CBP is a dependent variable, whereas the number of papers published in journals is an explanatory variable. We used funding variables to separate articles and their citations into two categories:

- o Funded articles refer to articles that are published in journals with funding.
- o Unfunded articles are defined as the published articles in journals that had not received any funding.

In order to test the power-law correlation between citations and total (funded + unfunded)/funded/unfunded articles, we used the model developed by Ronda-Pupo and Katz, (2016a, b). This model analyzes the correlation between Citation-based Performance (CBP) and the number of articles published in journals of computer science, medicine and economics,

$$\text{CBP} = kc^\alpha \quad (1)$$

CBP = for citation-based performance.

$c$  = the number of articles (total: funded + unfunded/funded/unfunded).

$k$  = constant.

$\alpha$  = the scaling factor.

Parameter  $\alpha$  is a measure of the magnitude of the Matthew effect (Ronda-Pupo, 2016; Katz, 1999) and can be used to explain the exponent of the scaling correlation between articles that are published in journals and their citations. This parameter ( $\alpha$ ) and the parameter of the correlation  $k$  were calculated with the Ordinary Least Squares (OLS) method (Leguendre & Leguendre, 2012). According to Ronda-Pupo and Katz, (2017), since the Web of Science data is relatively noise-free, and as there is minor error in measuring the number of citations of articles, OLS was used as a line fitting method. The scaling exponent  $\alpha$  can be used to predict the behavior of correlation.

In particular,

- When  $\alpha > 1$ , then CBP increases non-linearly, and citations of articles grow faster than the number of articles published in computer science, medicine, and economics jour-

nals. Thus, correlations are non-linear, and we have a positive Matthew effect or positive cumulative advantage.

- If  $\alpha < 1$ , the correlation is sub-linear, and its magnitude is a measure of the inverse Matthew effect or cumulative disadvantage. It means that the citations of articles grow slower than the number of published papers in journals.
- Finally, when  $\alpha = 1$ , the correlation is linear, and citations and articles published in journals grow at the same rate (Ronda-Pupo and Katz, 2017).

Student’s *T*-statistics were used to confirm if the scaling exponents of the power-law correlation suggested is significant.

## Results and discussion

Table 1 shows documents retrieved from the Web of Science database, organized per funding status.

We analyzed 55,071 documents, including articles and reviews, with 870,807 citations published in 712 journals of *computer science*. 34,080 (61.88%) of documents had received funding, and 70.06% of all citations are associated with this category of documents. Furthermore, 38.12% of the papers published in computer science were unfunded, and 29.94% of the citations are associated with this category. We did not consider conference proceedings as they are rarely funded. In brief, the main patterns in this discipline suggest that computer science journals published more funded papers than unfunded ones.

We also analyzed 113,310 documents with 1,967,600 citations in 973 *medicine* journals. In this field, 71.13% of the citations account for 50.14% of all documents, which are funded research. Medicine journals published equally funded and unfunded documents.

Results also show 23,398 documents in 537 *economics* journals in which 51% of citations accounted for 38.98% of all documents, which were funded papers. Economics journals published more unfunded papers than funded.

In general, funded documents received more citations than unfunded papers in all three fields. These findings suggest that computer science, medicine, and economics have a high correlation between funding for research and studies being cited at least once. These

**Table 1** The number of documents and citations of computer science, medicine, and economics according to funding status

Disciplines	Sources	Number of papers	%	Citations	%
Computer science	Funded	34,080	61.88	610,100	70.06
	Unfunded	20,991	38.12	260,707	29.94
	Total (funded + unfunded)	55,071	100	870,807	100
Medicine	Funded	56,810	50.14	1,399,477	71.13
	Unfunded	56,500	49.86	568,123	28.88
	Total (funded + unfunded)	113,310	100	1,967,600	100
Economics	Funded	9122	38.98	145,773	51
	Unfunded	14,276	61.02	145,617	49
	Total (funded + unfunded)	23,398	100	291,390	100

findings confirm results by Wang and Shapira's study (2015) about the impact of research sponsorship on the field of nanotechnology science research.

The power-law behavior of the distributions of citations and the number of papers of journals was analyzed using the procedure developed by Clauset et al. (2009). We summarize the procedure in two steps:

1. estimating parameters  $X_{min}$  and  $\alpha$ ;
2. comparing the power-law with alternative distributions. We also analyzed the power-law correlation between the number of citations and the number of papers of journals.

### **Verification of the power-law distribution for the papers in computer science, medicine, and economics**

Table 2 shows results of fitting data to a power-law distribution by 2500 iterations using the Monte-Carlo bootstrapping analysis to distribute citations to the following three data sets. Each of the datasets includes total (funded + unfunded), funded, and unfunded papers. Statistical evidence seems, in general, to support the empirical results of the power-law distribution for all three categories of papers in computer science, medicine, and economics. Interestingly, in Total categories, medicine has the lowest  $X_{min}$ , showing that with the lowest number of primary citations, the Matthew effect begins, and then citations start scaling exponentially. Economics has the highest  $X_{min}$ : journals need more initial citations before their exponent scaling starts.

### **Comparing the power-law with alternative distributions**

According to Clauset et al. (2009), we cannot conclude that if the  $p$ -value of the power-law distribution is significant ( $p > 0.1$ ), then it would be the best fit for the data. We compared the power-law model to lognormal, exponential, Poisson, stretched exponential, and power-law with exponential cut-off distributions as possible alternatives. According to Clauset et al. (2009), by using log-likelihood ratios (LR), we can identify which distribution can be considered as the best option. A distribution with the most negative LR is assumed to be the best distribution to fit the data. If the LR values of alternative distributions are positive, we can conclude that the power-law model is a plausible fit for the data. Table 3 shows the comparison between different distributions, including the magnitude of the scaling exponent  $\alpha$ , its  $p$ -value, and the log-likelihood ratio test (LR) for alternative distributions. The final column of Table 3 shows the statistical support indicating if the power-law fits for distributions. We used 'Moderate' to indicate that the power-law is a good fit, but there are other plausible alternatives as well. In some cases, the notion of 'with cut-off' indicates that the power-law with exponential cut-off is clearly preferable over the pure power-law (cf. Clauset et al. 2009, p. 687).

*Computer Science:* The  $p$ -value of the power-law distribution for computer science in the total (funded + unfunded), funded, and unfunded categories was significant (0.36, 0.7, 0.79), and we can conclude that the power-law distribution can be a plausible fit for our data. In the total (funded + unfunded) category, the  $p$ -values for exponential and Poisson distributions were statistically significant ( $p < 0.1$ ). Therefore, the exponential and Poisson distributions cannot be completely ruled out for our data. Still, the power-law distribution was likely a good fit as the LR of the alternatives were positive. In the funded sub-dataset, the  $p$ -values of exponential, Poisson, and stretched exponential distributions were



**Table 2** Results of fitting the power-law to the datasets

	Computer science			Medicine			Economics		
	Total (funded + unfunded)	Funded	Unfunded	Total (funded + unfunded)	Funded	Unfunded	Total (funded + unfunded)	Funded	Unfunded
	Xmin	69	72	57	57	51	36	74	47
$\alpha$	2.89***	2.89***	2.86***	2.54***	2.45***	3.05***	3.12***	3.07***	2.83***
KS	0.013	0.013	0.014	0.007	0.007	0.010	0.016	0.020	0.020

\*\*\* denotes significance when Power-Law  $p$ -value  $> 0.1$ . According to Clauset et al. (2009), Xmin is the value when the power-law begins.  $\alpha$  is the scaling factor (slope of the log–log regression line). The Kolmogorov–Smirnov (KS) statistic quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution

**Table 3** Results of comparing the power-law with alternative distributions

Fields	Funding	Power-law		Lognormal		Exponential		Poisson		Stretched exponential		Power-law + cut-off		Support for power-law	
		p	LR	p	LR	p	LR	p	LR	p	LR	p	LR	p	LR
Computer science	Total*	0.36	1.09	0.27	3.30	0.001	2.46	0.01	13.17	1.22	0.00	0.99	Moderate		
	funded	0.70	1.15	0.24	3.11	0.001	3.42	0.001	2.25	0.02	0.00	0.99	Moderate		
	unfunded	0.79	-0.29	0.76	3.81	0.001	0.00	2.23	0.62	0.53	-0.5	0.43	Moderate		
Medicine	Total*	0.24	-0.5	0.55	11.76	0.001	10.76	0.001	2.49	0.01	-1.48	0.04	with cut-off		
	funded	0.12	-0.83	0.40	13.18	0.001	11.28	0.001	2.54	0.01	-1.83	0.01	with cut-off		
	unfunded	0.45	-0.39	0.69	7.67	0.001	9.26	0.001	1.40	0.15	-0.73	0.33	Moderate		
Economics	Total*	0.74	-0.10	0.91	3.44	0.001	4.58	0.001	0.69	0.48	-0.31	0.71	Moderate		
	funded	0.41	-0.16	0.86	3.36	0.001	4.05	0.001	0.64	0.52	-0.13	0.87	Moderate		
	unfunded	0.34	-1.17	0.23	3.12	0.001	4.70	0.001	-0.98	0.33	-1.23	0.03	with cut-off		

\*Moderate\* indicates that the power-law is a good fit, but there are other plausible alternatives as well. 'with cut-off' indicates that the power-law with exponential cut-off is clearly preferable over the pure power-law (Clauset et al. 2009, p.687). \* Total = funded + unfunded

significant (0.001, 0.001, 0.02), so they could not be completely dismissed merely because their LR was 3.11, 3.42, and 2.25. Moreover, in the unfunded sub-dataset, the  $p$ -value of the exponential distribution was also statistically significant, but the LR was 3.81. These results suggest that we cannot rule out the exponential distribution for this data. The power-law seems to have a good fit for the distribution of citations.

*Medicine:* All alternative distributions had significant  $p$ -values for total (funded + unfunded) and funded data. The power-law distribution with cut-off had the most negative value of LR (− 1.48 for the total (funded+ unfunded) category, − 1.83 for the funded category), so this distribution was preferable compared to the pure power-law distribution. In unfunded data, the  $p$ -values of the exponential and Poisson are also statistically significant with  $p < 0.1$ . Accordingly, the exponential and Poisson distributions could not be completely ruled out in this category. However, the power-law distribution was marked as ‘Moderate’ as it was favorable because of the positive LR of alternatives. In particular, the power-law distribution with cut-off was generated from the pure one, and Katz, (2016) asserts, "models show that the probability distribution tends to evolve from a power-law with an exponential cut-off to a pure power-law given enough time". In life sciences, Su and Hogenesch, (2007) show power-law-like distributions in gene annotation (measuring links to the biomedical literature) and research funding, measured by gene references in funded grants (cf. Lewison & Dawson, 1998).

*Economics:* In the journals of this field, for total (funded + unfunded) and funded categories, the  $p$ -values for exponential and Poisson distributions were significant ( $p < 0.1$ ). Therefore, the exponential and Poisson distributions could not be completely ruled out for our data in these two classes. Still, the power-law distribution was a plausible fit as the LR values of other distributions were positive. In the unfunded sub-dataset, the statistical  $p$ -values of exponential, Poisson, and power-law with cut-off distributions were significant (0.001, 0.001, 0.03). This distribution was preferable compared to the pure power-law distribution in the presence of the most negative LR value of power-law with cut-off. The results seem to support that the power-law can be a good fit in all categories of our dataset.

**Power-law correlation between citation-based performance and number of papers in three categories of total (funded + unfunded), funded, and unfunded documents**

Table 4 shows that all correlations are statistically significant,  $p$ -value  $< 0.001$ .

In the field of *computer science*, Table 4 reveals that the power-law correlation between the number of papers and citations is statistically significant  $t(1,707)$ , 10.6,  $p < 0.001$ . The power-law correlation between funded papers and citation-based performance is

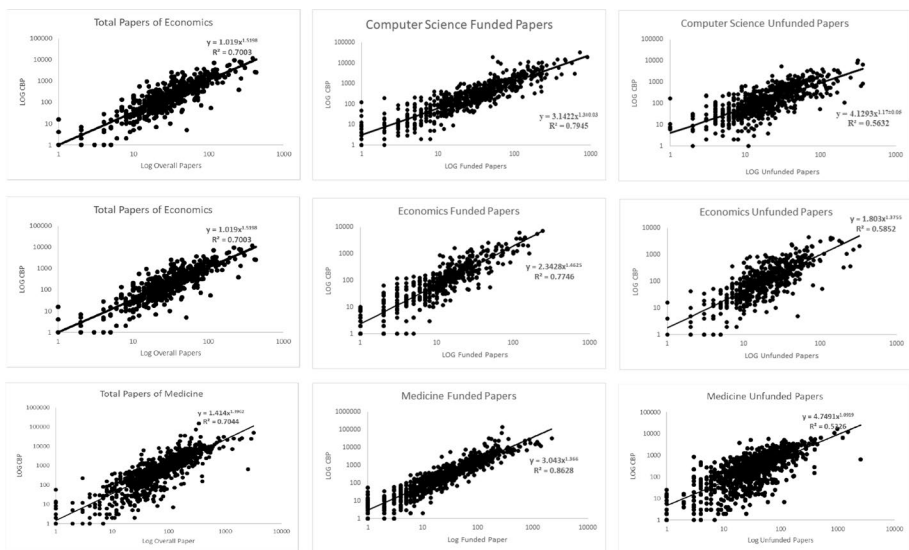
**Table 4** Values of the exponents for the power-law correlation

	Computer science			Medicine			Economics		
	Total	Funded	Unfunded	Total	Funded	Unfunded	Total	Funded	Unfunded
$\alpha$	1.31***	1.30***	1.17***	1.39***	1.36***	1.09***	1.51***	1.46***	1.37***
(SD)	(0.05)	(0.03)	(0.05)	(0.02)	(0.01)	(0.01)	(0.04)	(0.03)	(0.04)
$R^2$	0.69	0.79	0.56	0.71	0.86	0.53	0.70	0.77	0.58
N	707	660	705	972	866	963	537	470	534

Significance: \*\*\*,  $p$ -value  $< 0.001$ ; Total = funded + unfunded

statistically significant:  $t(1, 660)$ , 9.3,  $p < 0.001$ . The power-law correlation between unfunded documents and the number of citations is statistically significant:  $t(1, 705)$ , 11.7,  $p < 0.001$  (cf. Figure 1). The results suggest that the Matthew effect (Merton, 1968, 1988) is greater for funded articles compared with unfunded documents. The estimation of  $\alpha$  for funded papers is  $1.30 \pm 0.03$ , which means that the citation performance of research will increase by  $2^{1.30}$  or 2.46 times when they double the number of papers. The exponent  $> 1$  shows that citations grow faster than the rate of publication for funded papers. The exponent for the relationship between citations and unfunded papers is  $1.17 \pm 0.05$ . To put it differently, for a doubling of numbers of unfunded publications, the number of citations increases  $2^{1.17}$  or 2.25 times. Although there is a positive exponential relationship, the alpha is low compared to funded category. With a higher magnitude, doubling the number of funded papers increases the citation performance of research. The higher exponent for the funded papers suggests that an approach of increasing funding is required in this field to have a greater impact on the diffusion of science in society.

In the field of *medicine*, the power-law correlation between papers in funded + unfunded and citations is statistically significant:  $t(1,972)$ , 9.02,  $p < 0.001$ . The power-law correlation between funded papers and citation-based performance is also statistically significant:  $t(1, 866)$ , 7.18,  $p < 0.001$ . Finally, the power-law correlation between unfunded documents and the number of citations is statistically significant:  $t(1, 963)$ , 14.9,  $p < 0.001$ . The  $\alpha$  value of the total (funded + unfunded) category is estimated at  $1.39 \pm 0.02$ , suggesting a positive Matthew effect of  $2^{1.39}$  or 2.62 between the citation impact and the number of papers published in journals. Figure 1 shows that the power-law correlation between citations and total (funded + unfunded), funded, and unfunded research in medicine has a greater exponential scaling relationship with an  $\alpha$  value of 1.36 in funded research compared to the unfunded category having an  $\alpha$  value of 1.09. With an increase in the number of funded papers, the scaling of the citation performance for funded research increases by about  $2^{1.36}$  or 2.57 times. The expected number of citations increases by  $2^{1.09}$  or 2.13 times when the



**Fig. 1** The power-law relationship between citation performance and papers in total (funded + unfunded), funded, and unfunded of computer science, medicine, and economics

number of unfunded papers published in medicine journals doubles. The Matthew effect is stronger for funded papers than unfunded papers. In fact, by increasing the number of funded papers in this field, we can expect a higher magnitude of citation impact compared to unfunded category.

In the field of *economics*, the power-law correlation between papers in funded+unfunded and citations is statistically significant  $t(1,537)$ , 10.1,  $p < 0.001$ . The power-law correlation between funded papers and citation-based performance is statistically significant  $t(1, 470)$ , 19.63,  $p < 0.001$ . The power-law correlation between unfunded documents and the number of citations is statistically significant  $t(1, 534)$ , 10.8,  $p < 0.001$ . The discipline of economics also shows a scaling relationship. The value of  $\alpha$  for this field in the total (funded+unfunded) set is estimated at  $1.51 \pm 0.04$ , which means the citation performance of research will increase by  $2^{1.51}$  or 2.85 times when they double the number of published papers. Figure 1 shows the power-law correlation results between citations and total (funded+unfunded), funded, and unfunded research. The value of  $\alpha$  in funded papers is  $1.46 \pm 0.03$ , indicating that by doubling the number of funded documents in this field, the number of citations multiplies by more than two times: about  $2^{1.46}$  or 2.75. The Matthew effect is greater than the unfunded dataset, in which the value of  $\alpha$  is estimated to be  $1.37 \pm 0.047$ .

Finally, Table 4 also shows that coefficients of  $R^2$  of the funded datasets in computer science, medicine, and economics are higher than  $R^2$  of the power-law model for unfunded papers. As a result, the model has more explanatory power for the funded dataset.

Overall, then, statistical evidence of this inductive study based on three vital disciplines can be generalized in the following basic properties of dynamics of science.

Property 1. Research citations of funded and unfunded papers in different disciplines evolve with a power-law distribution.

Property 2. The number of articles and citation-based performance in different disciplines have a scaling correlation.

Property 3. The Matthew effect is greater for funded articles compared to unfunded documents in different disciplines.

## Conclusions

The present study explored the scaling relationship between citation-based performance and articles published in vital disciplines given by: computer science, medicine and economics. We analyzed the distribution of citations for papers in total (funded+unfunded), funded and unfunded researches separately. The results suggested that the power-law with exponential cut-off distribution analysis is the best fit for distributing citations to total (funded+unfunded) and funded papers of the medicine and unfunded papers of economics journals. All three categories of computer science, unfunded papers of the field of medicine, total (funded+unfunded) and funded papers of the field of economics are supported with the power-law distribution. Table 3 shows that the power-law is a plausible distribution for these datasets.

In *computer science*, results suggest a scaling correlation between the number of articles published in journals and citation-based performance (CBP) as:  $CBP \approx C^{1.31}$  for the total (funded+unfunded) dataset,  $CBP \approx C^{1.30}$  for unfunded papers. Results reveal that the power-law correlation between the number of funded papers and citations in the field of

computer science is statistically significant. Findings here show that the Matthew effect in funded papers is more powerful in comparison to unfunded studies.

With regards to the field of *medicine*, results also show a significant scaling correlation between the number of articles and citation-based performance (CBP), as:  $CBP \approx C^{1.39}$  for the total (funded + unfunded) dataset,  $CBP \approx C^{1.36}$  for funded papers, and  $CBP \approx C^{1.09}$  for unfunded papers. Findings also show that the Matthew effect in funded papers is greater than that of the unfunded dataset.

Finally, with regards to the field of *economics*, our findings show a scaling relationship between the number of papers and citation-based-performance (CBP) as:  $CBP \approx C^{1.51}$  for the total (funded + unfunded) dataset,  $CBP \approx C^{1.46}$  for funded papers, and  $CBP \approx C^{1.37}$  for unfunded papers. The Matthew effect in funded papers is also more considerable in comparison to unfunded research.

These results have main implications of research policy because a strategy of increasing funding for research can generate a higher diffusion of scientific results and a higher impact of science in society (Coccia et al. 2015; Coccia, 2005a, b).

These conclusions are certainly tentative. There is a need for much more detailed research into the relations under study here. As a future development, additional empirical research can further clarify an optimal scheme of research funding between different disciplines in order to support a vast diffusion of vital discoveries and new technologies for progress and well-being in human society.

**Acknowledgements** We would like to thank Professor J. Sylvan Katz for his helpful instruction on power-law analysis.


## References

- Alstott, J., Bullmore, E., & Plenz, D. (2014). Powerlaw: A Python package for analysis of heavy-tailed distributions. *PLoS ONE*, 9(1), e85777. <https://doi.org/10.1371/journal.pone.0085777>
- Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512.
- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4), 661–703.
- Coccia, M. (2005a). A Scientometric model for the assessment of scientific research performance within public institutes. *Scientometrics*, 65(3), 307–321. <https://doi.org/10.1007/s11192-005-0276-1>
- Coccia, M. (2005b). A taxonomy of public research bodies: A systemic approach. *Prometheus*, 23(1), 63–82. <https://doi.org/10.1080/0810902042000331322>
- Coccia, M. (2018). General properties of the evolution of research fields: A scientometric study of human microbiome, evolutionary robotics and astrobiology. *Scientometrics*, 117(2), 1265–1283. <https://doi.org/10.1007/s11192-018-2902-8>
- Coccia, M. (2019). Why do nations produce science advances and new technology? *Technology in society*, 59, 101124. <https://doi.org/10.1016/j.techsoc.2019.03.007>
- Coccia, M. (2020). The evolution of scientific disciplines in applied sciences: Dynamics and empirical properties of experimental physics. *Scientometrics*, 124, 451–487. <https://doi.org/10.1007/s11192-020-03464-y>
- Coccia, M., & Bozeman, B. (2016). Allometric models to measure and analyze the evolution of international research collaboration. *Scientometrics*, 108(3), 1065–1084. <https://doi.org/10.1007/s11192-016-2027-x>
- Coccia, M., & Wang, L. (2016). Evolution and convergence of the patterns of international scientific collaboration. *Proceedings of the National Academy of Sciences of the United States of America*, 113(8), 2057–2061. <https://doi.org/10.1073/pnas.1510820113>
- Coccia, M., Falavigna, G., & Manello, A. (2015). The impact of hybrid public and market-oriented financing mechanisms on scientific portfolio and performances of public research labs: A scientometric analysis. *Scientometrics*, 102(1), 151–168. <https://doi.org/10.1007/s11192-014-1427-z>

- Cronin, B., & Shaw, D. (1999). Citation, funding acknowledgment and author nationality relationships in four information science journals. *Journal of Documentation*, 55(4), 402–408.
- de Solla Price, D. J. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27, 292–306.
- Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., et al. (2018). Science of science. *Science*. <https://doi.org/10.1126/science.aao0185>
- Gao, J. P., Su, C., Wang, H. Y., Zhai, L. H., & Pan, Y. T. (2019). Research fund evaluation based on academic publication output analysis: The case of Chinese research fund evaluation. *Scientometrics*, 119(2), 959–972.
- Gillespie, C. S. (2015). Fitting heavy tailed distributions: The powerLaw package. *Journal of Statistical Software*, 64, 1–16.
- Gök, A., Rigby, J., & Shapira, P. (2016). The impact of research funding on scientific outputs: Evidence from six smaller European countries. *Journal of the Association for Information Science and Technology*, 67(3), 715–730. <https://doi.org/10.1002/asi.23406>
- Hicks, D., & Katz, J. S. (2011). Equity and excellence in research funding. *Minerva*, 49(2), 137–151.
- Huang, Z., Chen, H., Li, X., & Roco, M. C. (2006). Connecting NSF funding to patent innovation in nanotechnology (2001–2004). *Journal of Nanoparticle Research*, 8(6), 859–879.
- Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. *Journal of Public Economics*, 95(9–10), 1168–1177.
- Katz, J. S. (2000). Scale-independent indicators and research evaluation. *Science and Public Policy*, 27(1), 23–36.
- Katz, J. S. (2005). Scale-independent bibliometric indicators. *Measurement: Interdisciplinary Research and Perspectives*, 3(1), 24–28. [https://doi.org/10.1207/s15366359mea0301\\_3](https://doi.org/10.1207/s15366359mea0301_3)
- Katz, J. S. (2016). What is a complex innovation system? *PLoS ONE*, 11(6), e0156150.
- Leguendre, P., & Leguendre, L. (2012). *Numerical ecology* (3rd ed.). Elsevier B. V.
- Lewis, G., & Dawson, G. (1998). The effect of funding on the outputs of biomedical research. *Scientometrics*, 41, 17–27. <https://doi.org/10.1007/BF02457963>
- Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Academy of Sciences*, 16(1), 317–323.
- Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *Isis*, 79(4), 606–623.
- Morillo, F. (2020). Is open access publication useful for all research fields? Presence of funding, collaboration and impact. *Scientometrics*, 125, 689–716. <https://doi.org/10.1007/s11192-020-03652-w>
- Pao, M. L. (1991). On the relationship of funding and research publications. *Scientometrics*, 20, 257–281. <https://doi.org/10.1007/BF02018158>
- Quinlan, K. M., Kane, M., & Trochim, W. M. K. (2008). Evaluation of large research initiatives: Outcomes, challenges, and methodological considerations. *New Directions for Evaluation*, 118, 61–72. <https://doi.org/10.1002/ev.261>
- Rigby, J. (2013). Looking for the impact of peer review: Does count of funding acknowledgements really predict research impact? *Scientometrics*, 94(1), 57–73.
- Ronda-Pupo, G. A. (2017). The citation-based impact of complex innovation systems scales with the size of the system. *Scientometrics*, 112(1), 141–151.
- Ronda-Pupo, G. A., & Katz, J. S. (2016a). The scaling relationship between citation-based performance and coauthorship patterns in natural sciences. *Journal of the Association for Information Science and Technology*, 68(5), 1257–1265.
- Ronda-Pupo, G. A., & Katz, J. S. (2016b). The power-law relationship between citation-based performance and collaboration in articles in management journals: A scale-independent approach. *Journal of the Association for Information Science and Technology*, 67(10), 2565–2572.
- Ronda-Pupo, G. A., & Katz, J. S. (2018). The power law relationship between citation impact and multi-authorship patterns in articles in Information Science & Library Science journals. *Scientometrics*, 114(3), 919–932.
- Rousseau, R (2000, September). Are multi-authored articles cited more than single-authored ones? Are collaborations with authors from other countries more cited than collaborations within the country? A case study. In Proceedings of the second Berlin workshop on scientometrics and informetrics. Collaboration in Science and Technology. Gesellschaft für Wissenschaftsforschung: Berlin (pp.173–176)
- Shen, C.-C., Hu, Y.-H., Lin, W.-C., Tsai, C.-F., & Ke, S.-W. (2016). Research impact of general and funded papers. *Online Information Review*, 40(4), 472–480. <https://doi.org/10.1108/OIR-08-2015-0249>
- Su, A. I., & Hogenesch, J. B. (2007). Power-law-like distributions in biomedical publications and research funding. *Genome Biology*, 8(4), 404. <https://doi.org/10.1186/gb-2007-8-4-404>

- Van Raan, A. (1998). The influence of international collaboration on the impact of research results: Some simple mathematical considerations concerning the role of self-citations. *Scientometrics*, 42(3), 423–428.
- Van Raan, A. F. J. (2006). Statistical properties of bibliometric indicators: Research group indicator distributions and correlations. *Journal of the American Society for Information Science and Technology*, 57(3), 408–430. <https://doi.org/10.1002/asi.20284>
- Van Raan, A. F. V. (2008). Bibliometric statistical properties of the 100 largest European research universities: Prevalent scaling rules in the science system. *Journal of the American Society for Information Science and Technology*, 59(3), 461–475.
- Wang, J., & Shapira, P. (2011). Funding acknowledgement analysis: An enhanced tool to investigate research sponsorship impacts: The case of nanotechnology. *Scientometrics*, 87(3), 563–586.
- Wang, J., & Shapira, P. (2015). Is there a relationship between research sponsorship and publication impact? An analysis of funding acknowledgments in nanotechnology papers. *PLoS ONE*, 10(2), e0117727.
- Web of Science. (2021). Web of Science, Search in: Web of Science Core Collection, [http://apps.webofknowledge.com/WOS\\_GeneralSearch\\_input.do?product=WOS&SID=E51SYgaLwJn6kp2iz2G&search\\_mode=GeneralSearch](http://apps.webofknowledge.com/WOS_GeneralSearch_input.do?product=WOS&SID=E51SYgaLwJn6kp2iz2G&search_mode=GeneralSearch) (Accessed 24 March 2021)
- Yan, E., Wu, C., & Song, M. (2018). The funding factor: A cross-disciplinary examination of the association between research funding and citation impact. *Scientometrics*, 115(1), 369–384.
- Ye, F. Y., & Rousseau, R. (2008). The power law model and total career h-index sequences. *Journal of Informetrics*, 2(4), 288–297. <https://doi.org/10.1016/j.joi.2008.09.002>
- Zhao, S. X., Lou, W., Tan, A. M., & Yu, S. (2018). Do funded papers attract more usage? *Scientometrics*, 115(1), 153–168.

## Authors and Affiliations

Saeed Roshani<sup>1</sup>  · Mohammad-Reza Bagherylooieh<sup>1</sup> · Melika Mosleh<sup>2</sup> · Mario Coccia<sup>3</sup>

Mohammad-Reza Bagherylooieh  
bagheri@atu.ac.ir

Melika Mosleh  
mxm1219@alumni.bham.ac.uk

Mario Coccia  
mario.coccia@cnr.it

<sup>1</sup> Department of Industrial Management, Faculty of Management and Accounting, Allameh Tabataba'i University, Tehran, Iran

<sup>2</sup> Birmingham Business School, College of Social Sciences, University of Birmingham, Birmingham, UK

<sup>3</sup> CNR—National Research Council of Italy Collegio Carlo Alberto, Via Real Collegio, 30-10024 Moncalieri, Torino, Italy