

Assessing the effect of the United States’ “citation advantage” on other countries’ scientific impact as measured in the Web of Science (WoS) database

Yves Gingras¹ · Mahdi Khelifaoui¹

Received: 5 October 2017 / Published online: 4 December 2017
© Akadémiai Kiadó, Budapest, Hungary 2017

Abstract Eugene Garfield, always insisted that citation analysis “can be used wisely or abused” and that it is “up to the scientific community to prevent abuse of the SCI by devoting the necessary attention to its proper and judicious exploitation” (Garfield in *Nat* 227:669–671, 1970). Dedicated to his memory, this paper aim to assess the significance of a parameter that is seldom taken into account in evaluation studies: the existence of a USA comparative citation (visibility) advantage *built in the database* and thus affecting countries that collaborate more with the USA than with other countries. We analyze how this USA citation advantage affects the measure of the scientific impact (usually measured through citations received) of major countries. The main conclusion coming out of this study is that, given the strong presence of the USA in the WoS database, the comparative rankings tend, by construction, to give a citation advantage to countries having the closest relation to that country.

Keywords Evaluation · Database · Biased indicators · Bibliometrics · Rankings

Introduction

Although not new (Inhaber and Alvo 1978; Rabkin and Inhaber 1979), the use of bibliometric databases to evaluate and compare the scientific production of countries, geographic regions or even institutions has now become a widespread practice in academia (for recent examples, see Bornmann and Leydesdorff 2013; Gul et al. 2015; Jurajda et al. 2017; Öquist and Benner 2015; Rodriguez-Navarro and Narin 2017) as well as among science policy makers (European Commission 2014; National Science Board 2016). Not surprisingly, the first studies to compare and rank the «top nations» in terms of scientific

✉ Yves Gingras
gingras.yves@uqam.ca

¹ Observatoire des sciences et des technologies, Centre interuniversitaire de recherche sur la science et la technologie, Université du Québec à Montréal, Montréal, Québec, Canada

productivity and impact underscored the dominance of the USA, the leading European Union countries and Japan (King 2004; May 1997). Other studies found that the USA globally outperformed other nations (Leydesdorff and Wagner 2009; Leydesdorff et al. 2014), while observing that it was gradually declining and losing ground, especially in terms of world share of publications, due to the rise of China and other Asian countries (Glänzel et al. 2008; Leydesdorff and Zhou 2005; Zhou and Leydesdorff 2006). Interest in evaluating and ranking has not only been limited to the most advanced nations but has also been applied to developing countries or regions (Collazo-Reyes 2014; Confraria and Godinho 2015; Gonzalez-Brambila et al. 2016; Sarwar and Hassan 2015). Various types of indicators have been used to evaluate national scientific performance and productivity of nations, some based on publications, like the ratio of publications to R&D spending (or to GDP), and others using citations, like the share of top 1% cited papers, or relative citation rates (Bar-Ilan 2008, pp. 27–28).

In this paper, we aim to assess the significance of a parameter that is seldom taken into account in this type of studies: the existence of a USA comparative citation (visibility) advantage *built in the database* and thus affecting countries that collaborate more with the USA than with other countries. We will analyze how this USA citation advantage affects the measure of the scientific impact (usually measured through citations received) of major countries. Considering this parameter is particularly relevant in light of the important weight of USA's publications and journals in the most commonly used bibliometric databases. For example, when comparing USA and UK natural sciences and engineering (NSE) journal coverage rates between the Web of Science (WoS) and the more comprehensive database Ulrich, Archambault et al. (2006) found an overrepresentation of USA and UK journals in the WoS of 19 and 36% respectively. Van Leeuwen et al. (2001) concluded that science impact indicators were very sensitive to the inclusion or not of other languages than English publications and that comparative evaluations of national research systems must take this factor into account. More recently, Mongeon and Paul-Hus (2016) have also pointed to biases in Scopus and the WoS induced by the overrepresentation of certain discipline categories and English-language journals, thus accentuating the weight of Anglo-Saxon countries in these databases.

Apart from its intrinsic scientific qualities, related to the epistemic stakes of its disciplinary field, a scientific paper's impact is also influenced by external factors such as the journal's impact factor (Larivière and Gingras 2010; Lozano et al. 2012), the number of coauthors (de Beaver 2004; Frenken et al. 2005; Hsu and Huang 2011; Larivière et al. 2015) and international collaboration (Aksnes 2003; Bordons et al. 1996; Glänzel 2001; Narin et al. 1991). Patterns of international collaboration between countries are themselves influenced by political, geographic, economic, cultural and linguistic affinities which usually arise from a shared historical past (Davidson Frame and Carpenter 1979; Katz 1994; Luukkonen et al. 1992; Zitt et al. 2000, Larivière et al. 2006). Also, the tendency to collaborate with a country rather than another has an effect on scientific impact (Lancho-Barrantes et al. 2013) and the gain in scientific impact rises when collaborating with high-impact countries (Guerrero-Bote et al. 2013).

Given the strong presence of USA's papers in the WoS and the importance of international collaboration in accruing citations, it makes sense to investigate how a country's scientific impact is influenced by its level of collaboration with the USA. More precisely, in this paper we aim, first, to assess how the propensity of a country to receive citations from USA publications influences its scientific impact, as measured by a normalized measure of citations. We also want to determine the strength of the relationship between the intensity of international collaboration with the USA and the contribution to citations

resulting from USA citations. If a strong relationship is demonstrated, this would mean that countries that collaborate more with the USA are more likely to attract citations from USA publications and thus increase their scientific impact comparatively to countries whose share of world publications might be equivalent but who collaborate less with the USA.

Given that the uses of research evaluations based on citation measures can affect the career of researchers as well as the very existence of whole university departments, this exercise is more than academic, for it raises the ethical question of the fairness of evaluation that is too often left implicit in using databases, be it Google Scholar, Scopus, WoS or other tailor-made databases like those of private firms like Academic Analytics (Basken 2016). In other words, it is important to closely analyze the effects of the content of a database on the chosen indicators in order to understand how that indicator behaves when comparing different countries. Not doing so would imply that policy conclusions could in fact be based on *implicit* decisions that collaborating with a given country is “better” independently of the national policy priorities in terms of scientific development. Databases are rarely neutral and evaluators should do their best to make sure they really measure what they want to measure, that is *scientific* impact and not simply historical and political relations between countries.

A recent example showing how an unreflexive use of an indicator could lead to bad policy decisions is provided by the latest *Science & Engineering Indicators* Report from the National Science Foundation. Using data from the Scopus database, in 2013, it shows that the USA and China had almost the same share of world publications, with 18.8 and 18.2% respectively (National Science Board 2016, Ch. 5, p. 92; Witze 2016). This may come as a surprise and could suggest that the USA should spike their R&D investments to keep their first place. But before reaching that conclusion, one should notice that this number is based on a fractional count of the papers attributed to each country, and this affects the meaning attributed to the numbers. But fractional counting is of course affected by the number of countries with which one collaborates. Since this number has risen much more for the USA than for China, it is mechanical that fractional counting will play against the USA. So if one considers international collaboration important, the number of collaborating countries is the important indicator even though it diminishes the contribution based on fractional counting of papers. Hence the average number of countries per paper for USA moved from 1.3 in 2001 to nearly 1.6 in 2013 while that for China started also at 1.3 in 2001 but rose only to 1.36 in 2013. Even the data source affects the results, for using WoS instead of Scopus and using the fractional method of contributions to papers, we get 21% for the USA and 14% for China in 2013. The difference between the two sets of numbers can easily be explained by the fact that Scopus contains more low-end journals than WoS and thus includes more papers from China, which appear in that long tail of less cited journals.

As this paper is dedicated to Eugene Garfield, we can only recall that he has always insisted that citation analysis “can be used wisely or abused” and that it is “up to the scientific community to prevent abuse of the SCI by devoting the necessary attention to its proper and judicious exploitation” (Garfield 1970). Thus, this paper aims to contribute to a better exploitation of the meaning of “impact”, based on citations, when that “impact” may be affected by variables other than the objective scientific content of the cited papers.

Methods

We considered all the papers published in 2010 by five major producers of science (USA, $N = 289,099$; Germany, $N = 82,867$; England, $N = 66,177$; France, $N = 60,382$; and Canada, $N = 46,913$) in the eight disciplines constituting the NSE category in the WoS (physics, chemistry, biology, mathematics, clinical medicine, biomedical research, Earth and atmosphere sciences and engineering). We defined a country's publication as a publication in which the address of at least one coauthor was located in that country.

Initially, we calculated the average of citations received by each of the 5 countries in each of the eight disciplines, over a period of 5 years, i.e. between 2011 and 2015. Then, we calculated the weighted average of citations for the world's NSE publications ($N = 1,075,606$), a measure that we used to calculate the Average of Relative Citations (ARC) as a unit of normalization for each of the 5 countries. The ARC received in the field of NSE over a period of 5 years was calculated as being the sum of the weighted average of citations of the 8 disciplines constituting the NSE field, $WAC[NSE]_{Country}$, divided by the weighted average of citations of the world's publications in NSE, $WAC[NSE]_{World}$. Whether for a given country or for the world, the value of $WAC[NSE]$ was determined by the following equation:

$$WAC[NSE] = \sum_{i=1}^8 \frac{Tot_cit_i}{Tot_art_i} \times \frac{Tot_art_i}{Tot_art_{NSE}} = \sum_{i=1}^8 \frac{Tot_cit_i}{Tot_art_{NSE}}$$

where i is the discipline; Tot_cit_i is the total number of citations received by a country (or the world) in this discipline, Tot_art_i is the total number of articles published by a country (or the world) in this discipline, and Tot_art_{NSE} is the world's total number of articles published in NSE.

Also, three categories of publications were considered: (1) all papers; (2) papers written in international collaboration (IC) and (3) papers written without IC, i.e. with all authors from the same country. For each category of publications, an ARC was calculated using the following equation:

$$ARC[NSE]_{Country} = \frac{WAC[NSE]_{Country}}{WAC[NSE]_{World}}$$

In a second step, we performed the same ARC calculation, this time subtracting the number of USA citations from the total count of citations received by each category of papers, in each discipline of the NSE field. Here, we defined a USA citation as coming from a publication in which at least one address was located in the USA. Using this definition means that citations from publications resulting from collaborations between the USA and all other countries were not counted either. The new value of $WAC[NSE]$ was obtained from the following equation:

$$WAC[NSE] = \sum_{i=1}^8 \frac{Tot_cit_i - USA_cit_i}{Tot_art_{NSE}}$$

where USA_cit_i is the number of citations received from USA publications by a given country (or the world) in a given discipline.

The difference between the two values of $WAC[NSE]$ measures the drop in scientific visibility (citations) of a group of publications, induced by not taking into account USA citations in the WoS database. We performed the same calculation for each of the four

other chosen countries, measuring each time the magnitude of the effect on the other countries' visibility.

In order to evaluate the relative visibility of a country's publications in the USA, we used another indicator, defined as the ratio of two variables: the percentage of citations given by the USA to this country's publications and the share of this country's world publications in the same discipline. This relative visibility factor is given by the following equation:

$$VIS_{USA \rightarrow Country} = \frac{N_{\text{cit}_{USA \rightarrow Country}}}{N_{\text{cit}_{USA \rightarrow World}}} \bigg/ \frac{N_{\text{pub}_{Country}}}{N_{\text{pub}_{World}}}$$

where $N_{\text{cit}_{USA \rightarrow Country}}$ is the number of citations given by the USA to another country, $N_{\text{cit}_{USA \rightarrow World}}$ is the number of all US citations, $N_{\text{pub}_{Country}}$ is the number of publications from the country and $N_{\text{pub}_{World}}$ is the number of world publications. We calculated this relative visibility factor for the eight NSE disciplines and for two categories of publications: those written in international collaboration (IC) and national ones. If $VIS_{USA \rightarrow Country}$ is equal to 1.0, this means that the proportion of citations received by the country from the USA is equivalent to its contribution to the world pool of publications. If $VIS_{USA \rightarrow Country}$ is superior (or inferior) to 1.0, this means that the country's publications is more (or less) visible in the USA than their real weight in terms of world share of publications would suggest, all other things being equal.

Another method of evaluating the relative visibility of a country's publications in the USA consisted in calculating the average percentage of this country's publications in all the references present in USA NSE publications, between 2010 and 2015 inclusively. The relative visibility factor was obtained by dividing the percentage of the country's publications in USA references by the share of the country's NSE world publications. While the previous measure, $VIS_{USA \rightarrow Country}$, was based on the count of citations received by papers published solely in 2010, this measure differed by the fact that a country's publication present as a reference in a USA publication could be dated from any year before 2015. The relative visibility factor was calculated, for both domestic and international USA publications, using the following equation:

$$Ref_{Country \rightarrow USA} = \frac{N_{\text{ref}_{Country \rightarrow Pub_US}}}{N_{\text{ref}_{World \rightarrow Pub_US}}} \bigg/ \frac{N_{\text{pub}_{Country}}}{N_{\text{pub}_{World}}}$$

where $N_{\text{ref}_{Country \rightarrow Pub_US}}$ is the number of publications of a given country present in the references of USA's publications; $N_{\text{ref}_{World \rightarrow Pub_US}}$ is the total number of references present in USA's publications.

Finally, in order to determine if there is a direct relationship between the intensity of international collaboration with the USA and the amount of citations received from USA publications, we tested the degree to which the two variables were correlated. The first variable was, for a given country, the percentage of NSE papers written in international collaboration with the USA. The second variable was the drop in ARC for this country when citations stemming from US publications were excluded. The test consisted in verifying if the drop in ARC was more significant when the percentage of collaboration with the USA was higher. We tested this hypothesis on papers published in two different years, 2010 and 1980, and cited during the 5 years following, using a list of 16 countries in the world top science producers: England, Germany, France, Canada, Italy, Spain, Sweden, Russia, Japan, China, South Korea, India, Turkey, Finland, Australia and Brazil.

Results

Figure 1 shows the results of ARC drop (ΔARC in absolute value) for the USA, England, Germany, Canada, France, and World NSE publications, following the subtraction of each of the five countries' contributions to citations from the calculation of the weighted average of citations. Comparatively to the other 4 countries, subtracting USA citations has clearly the most important effect (red hexagon in Fig. 1). England and Canada are the most affected from the exclusion of citations from the USA with a ΔARC of 0.51 and 0.47, respectively. They are followed by Germany (0.40) and France (0.36). It is not surprising to observe that Canada and England are more affected than Germany and France, given their geographic (for Canada), and cultural and linguistic proximity to the USA (for England). At the disciplinary level, this effect is more pronounced in biomedical research ($\Delta\text{ARC}_{\text{Eng}} = 0.58$, $\Delta\text{ARC}_{\text{Can}} = 0.53$, $\Delta\text{ARC}_{\text{Ger}} = 0.48$, $\Delta\text{ARC}_{\text{Fra}} = 0.43$) and in clinical medicine, while it is much less pronounced in chemistry ($\Delta\text{ARC}_{\text{Eng}} = 0.29$, $\Delta\text{ARC}_{\text{Can}} = 0.22$, $\Delta\text{ARC}_{\text{Ger}} = 0.21$, $\Delta\text{ARC}_{\text{Fra}} = 0.17$) and mathematics.

Given the frequent use of “country rankings” based on citations, it is important to observe that excluding citations from the USA modifies the countries ranking order according to their ARC. For all NSE papers, when USA citations are counted, Canada (ARC = 1.357) ranks before Germany (ARC = 1.343) and France (ARC = 1.250). But when USA citations are not counted (for all countries), Germany (ARC = 0.947) ranks above France (ARC = 0.891), and France above Canada (ARC = 0.884). In both

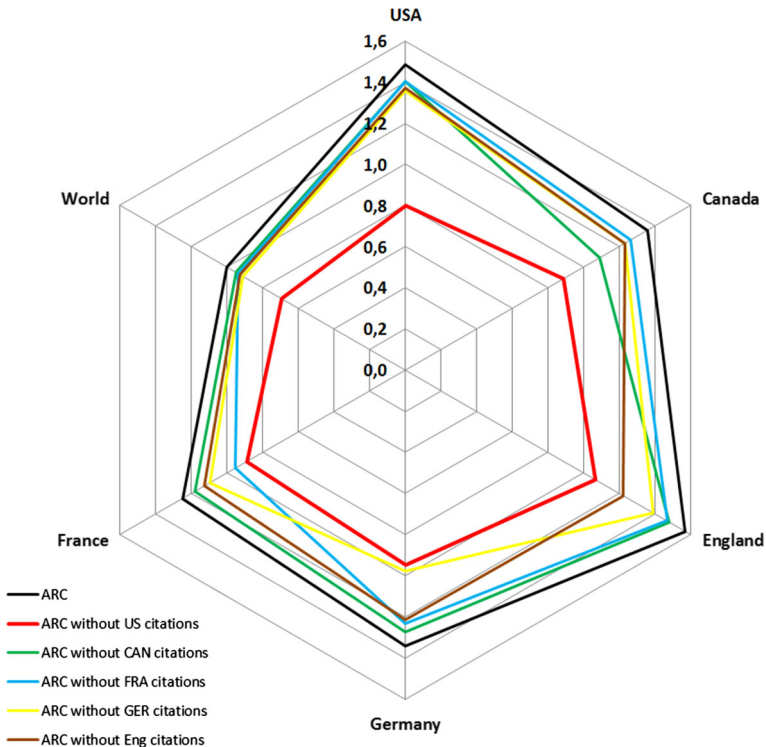


Fig. 1 ARC drop after subtracting US, English, Canadian, German or French citations (all NSE papers)

situations (counting or not counting US citations), England keeps its rank above the three other countries. Exactly the same change in ranking order is observed for NSE papers written in international collaboration. For papers written without international collaboration, Canada loses one place to Germany but still outranks France. For each of the eight disciplines, and for papers in and without international collaboration, which provides a total of 16 possible combinations, there are seven cases where Canada loses one or two positions to Germany and France after US citations are subtracted from the calculation of their ARC (Table 1).

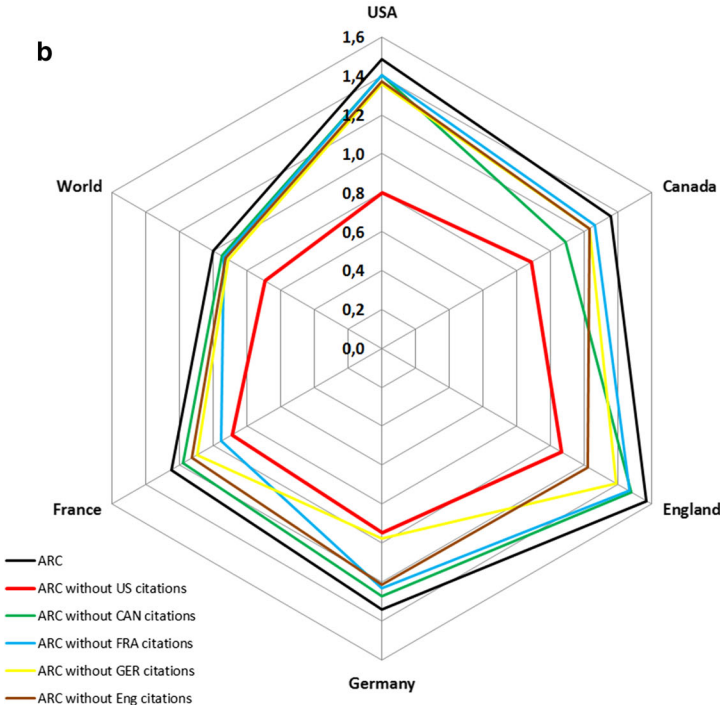
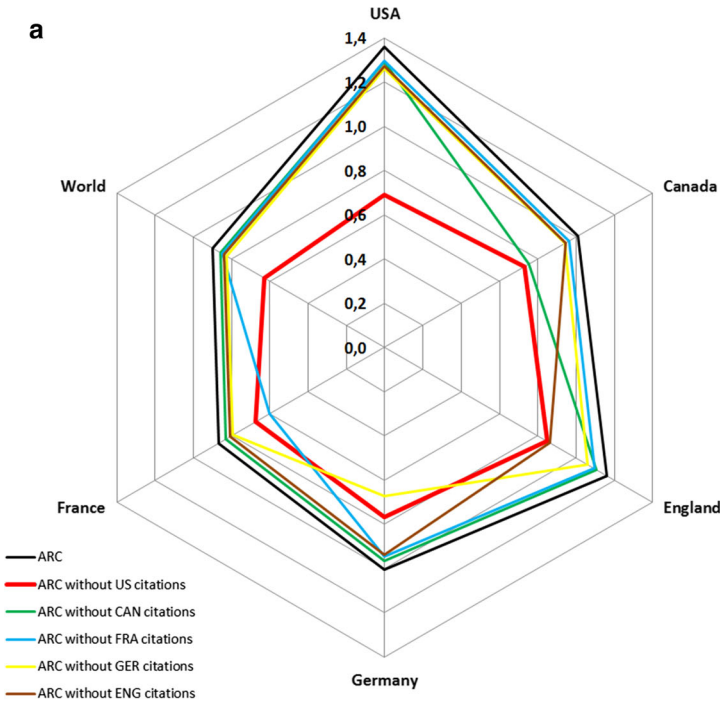
Figure 2a and b display the same calculation as Fig. 1, for papers written in international (Fig. 2a) and national (Fig. 2b) collaboration. For internationally co-authored papers, the ARC drop (after subtracting USA citations) is more accentuated than in Fig. 1, while the effect is less accentuated for national papers. This result shows that papers written in international collaboration are also usually more cited in the USA, which is not surprising considering that a good part (23% of all publications for Canada, 17% for England, 15% for Germany and 13.5% for France) of these papers are written with US coauthors, are mostly written in English and that, in general, papers written in international collaboration are more cited than national papers. From Figs. 1, 2a and b, we observe that, for England, Germany and France, the exclusion of USA citations has a more important effect on ARC drop than excluding citations from their own country. For example, $\Delta ARC_{Ger} = 0.40$ without USA citations, while $\Delta ARC_{Ger} = 0.37$ when excluding German citations. Similarly, $\Delta ARC_{Can} = 0.47$ without USA citations, while $\Delta ARC_{Can} = 0.27$ without Canadian citations. This result holds for the four countries in the three categories of papers defined above, except for two cases: Germany and France in the category of national papers (Fig. 2a).

Table 2 shows the values of the visibility factor calculated for all five countries. First, we observe that, as is now well known, for most disciplines, publications in international collaborations are generally more visible in the USA than strictly national publications, except for the case of USA domestic publications; interestingly, those are more visible in that country than their own publications in international collaborations. There is also a particular case in chemistry, where German, French and Canadian national publications have more visibility than their international publications in the USA. In general, depending on the discipline and the category of publication, Canada or England arrives in second position. For national papers, England’s visibility is superior or equivalent to Canada’s, in all disciplines. For international collaborations, Canada has more visibility than England in

Table 1 Ranking changes after subtracting USA citations

Discipline	Ranking with USA citations	Ranking without USA citations
Physics (WIC)	1. GER; 2. ENG; 3. CAN; 4. FRA	1. GER; 2. ENG; 3. FRA; 4. CAN
Physics (IC)	1. CAN; 2. ENG; 3. GER; 4. FRA	1. ENG; 2. GER; 3. CAN; 4. FRA
Biology (WIC)	1. ENG; 2. FRA; 3. CAN; 4. GER	1. ENG; 2. FRA; 3. GER; 4. CAN
Biomed. (WIC)	1. ENG; 2. GER; 3. CAN; 4. FRA	1. ENG; 2. GER; 3. FRA; 4. CAN
Biomed. (IC)	1. ENG; 2. CAN; 3. GER; 4. FRA	1. ENG; 2. FRA; 3. GER; 4. CAN
Medicine (IC)	1. FRA; 2. ENG; 3. CAN; 4. GER	1. FRA; 2. ENG; 3. GER; 4. CAN
Mathematics (IC)	1. ENG; 2. GER; 3. CAN; 4. FRA	1. ENG; 2. GER; 3. FRA; 4. CAN

WIC Without International Collaboration, IC International Collaboration



◀ **Fig. 2 a** ARC drop after subtracting US, English, Canadian, German or French citations (NSE papers in international collaboration). **b** ARC drop after subtracting US, English, Canadian, German or French citations (NSE papers without international collaboration)

Table 2 Visibility factor in the USA for the USA, England, Germany, Canada, and France

Discipline	VIS _{USA→USA}		VIS _{USA→ENG}		VIS _{USA→GER}		VIS _{USA→CAN}		VIS _{USA→FRA}	
	WIC	IC	WIC	IC	WIC	IC	WIC	IC	WIC	IC
All	2.50	1.63	1.16	1.47	0.88	1.24	1.05	1.53	0.72	1.15
Biology	2.63	1.71	1.83	1.51	0.89	1.20	1.21	1.61	0.99	1.18
Chemistry	3.45	2.12	1.35	1.59	1.15	1.00	1.37	0.95	0.93	0.78
Engineering	3.18	1.79	0.94	1.04	0.90	1.18	0.96	0.89	0.77	0.87
Mathematics	3.04	1.84	1.18	1.45	0.87	1.13	1.05	1.06	0.75	0.88
Clinical medicine	1.99	1.37	0.95	1.23	0.67	1.17	1.05	1.43	0.58	1.28
Physics	3.24	1.79	1.37	1.40	1.28	1.26	1.13	1.68	1.00	1.18
Biomed. research	1.99	1.43	1.10	1.38	0.92	1.18	0.86	1.47	0.73	1.14
Earth and space science	2.56	1.57	0.99	1.46	0.88	1.47	0.78	1.47	0.76	1.39

WIC Without International Collaboration, IC International Collaboration

four disciplines: biology, clinical medicine, biomedical research and physics. Germany is more visible than France in all disciplines and in both categories of publications, except in two cases: domestic publications in biology (0.89 for Germany vs 0.99 for France) and international publications in clinical medicine (1.17 for Germany vs 1.28 for France). France is the country which has more visibility factors inferior or equal to 1.0. It is the case for its domestic papers in all eight disciplines. In Table 3, we observe a “national citation bias” that makes national publications more visible in their country than international publications, in all disciplines and in both categories of publications. This bias is more

Table 3 National citation bias for the USA, England, Germany, Canada, and France

Discipline	VIS _{USA→USA}		VIS _{ENG→ENG}		VIS _{GER→GER}		VIS _{CAN→CAN}		VIS _{FRA→FRA}	
	WIC	IC	WIC	IC	WIC	IC	WIC	IC	WIC	IC
All	2.50	1.63	5.31	2.85	4.70	2.57	6.28	3.54	5.64	3.05
Biology	2.63	1.71	9.64	3.50	6.20	3.15	7.37	3.86	8.49	3.54
Chemistry	3.45	2.12	8.57	4.06	6.12	2.58	10.45	5.12	7.56	3.03
Engineering	3.18	1.79	7.35	3.26	6.50	3.50	7.31	3.65	8.18	3.45
Mathematics	3.04	1.84	9.04	4.54	6.98	3.37	8.55	4.35	7.46	3.34
Clinical medicine	1.99	1.37	3.83	2.27	3.51	2.40	5.06	2.81	4.14	3.20
Physics	3.24	1.79	7.74	3.38	5.85	2.32	10.02	5.46	6.91	2.82
Biomed. research	1.99	1.43	4.16	2.45	3.95	2.25	4.65	3.17	4.63	2.69
Earth and space science	2.56	1.57	4.95	2.42	4.46	2.34	5.81	3.27	5.51	2.74

WIC Without International Collaboration, IC International Collaboration

Table 4 Average relative presence of the USA, England, Germany, Canada and France in the references of US publications (2010–2015)

	Ref _{USA→USA}	Ref _{ENG→USA}	Ref _{GER→USA}	Ref _{CAN→USA}	Ref _{FRA→USA}
US pub. (WIC)	2.56	1.35	0.95	1.33	0.92
US pub. (IC)	2.21	1.83	1.36	1.56	1.36
World pub. (all)	1.72	1.57	1.23	1.34	1.23

WIC Without International Collaboration, IC International Collaboration

important for national papers, which means that these countries' national publications cite more papers from their own country, comparatively to their international publications.

Table 4 shows the values of visibility factors calculated for the five countries through their relative presence in the references of USA and world NSE publications, between 2010 and 2015. Similarly to the measure of visibility displayed in Table 1, and except for the USA, the four other countries' presence in the references of USA publications is more important when the USA publications are international. We also notice that England and Canada have a greater relative presence in USA references than Germany and France, for both categories of publications. The relative presence of England, Canada, Germany and France in the references of USA international publications is always superior to their relative presence in the world's total publications, while it is always inferior in the references of USA national publications. In this latter case, Germany (0.95) and France (0.92) are even less present than what would be expected given their world share of publications. The fact that the USA is relatively more present in the references of its own national publications (2.56) than in its international publications (2.21) is another demonstration of this country's national citation bias towards its own publications. This phenomenon is not limited to the USA but applies to the four other countries, in which it is even more accentuated (Table 5).

Figure 3 shows a strong correlation ($R^2 = 0.81$) between the variation of normalized citations (Δ ARC), calculated for all papers published in 2010 by a given country, and the percentage of papers written by this country in collaboration with the USA. This means that, at the aggregate level, *the more a country publishes papers in collaboration with the USA, the more likely it is to receive citations from USA publications*. Countries with a small share of their published papers in collaboration with the USA are the ones who have the smallest loss in visibility (Δ ARC). These are the Asian and emerging countries (Turkey, India, China, Russia, South Korea, Japan and Brazil). In the middle of the curve are the European countries (Spain, France, Italy, Germany), while Anglo-Saxon (England, Canada, Australia) and Nordic (Sweden, Finland) countries are at the top of the curve as they collaborate the most with the USA and are the ones that benefit the most from

Table 5 Average relative presence of the 5 countries in the references of their own publications (2010–2015)

Publication	Ref _{USA→USA}	Ref _{ENG→ENG}	Ref _{GER→GER}	Ref _{CAN→CAN}	Ref _{FRA→FRA}
(WIC)	2.56	4.48	3.74	5.05	4.58
(IC)	2.21	3.77	3.11	3.86	3.81

WIC Without International Collaboration, IC International Collaboration

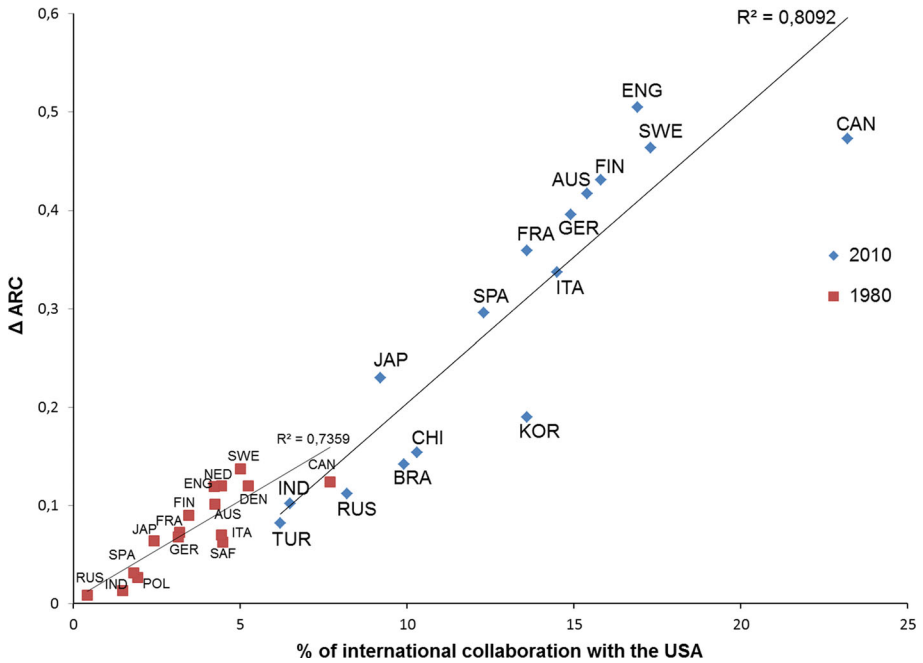


Fig. 3 Relationship between Δ ARC and the % of international collaboration with the USA (1980 and 2010)

citations from the USA. The second curve in Fig. 3 shows a less strong, however still important, correlation ($R^2 = 0.73$) between the variation of normalized citations (Δ ARC), calculated for all papers published in 1980, and the percentage of papers written in collaboration with the USA. Due to the small total publications of China, South Korea, Turkey and Brazil, referenced in the WoS in 1980, numbers which are not large enough to provide a reliable idea of the real percentage of collaboration of these countries with the USA, four other countries were used instead: Poland, Denmark, South Africa and The Netherlands. The percentage of international collaboration between the USA and other countries in general being much inferior in 1980 than in 2010, the loss in visibility (Δ ARC) is less important in the 1980 curve. This adds another confirmation that the gain in citations from the USA is strongly related to the amount of collaboration with this country.

As shown in Fig. 4a and b, the same correlation as in Fig. 3 is obtained when replacing the USA by France or China for international collaborations in 2010. In comparison to the USA, however, the correlation for France and China is less strong ($R^2 = 0.76$ for France and $R^2 = 0.59$ for China) and, in general, the loss in citations much smaller, which means that less collaboration with France or China has a lesser impact on the drop in ARC value than less collaboration with the USA.

Figure 5 shows the relationship between the percentage of USA collaboration with other countries and the percentage of citations made by the USA to these countries. The correlation between these two variables is strong ($R^2 = 0.83$), meaning that the more the USA collaborates with a country, the more it will cite publications from this country. Again, Asian and emerging countries, except China, are at the bottom of the curve, while European countries and Canada are at the top. China’s position on Fig. 5 is interesting. It is the country with which the USA collaborates the most (4.8%), but at the same time it

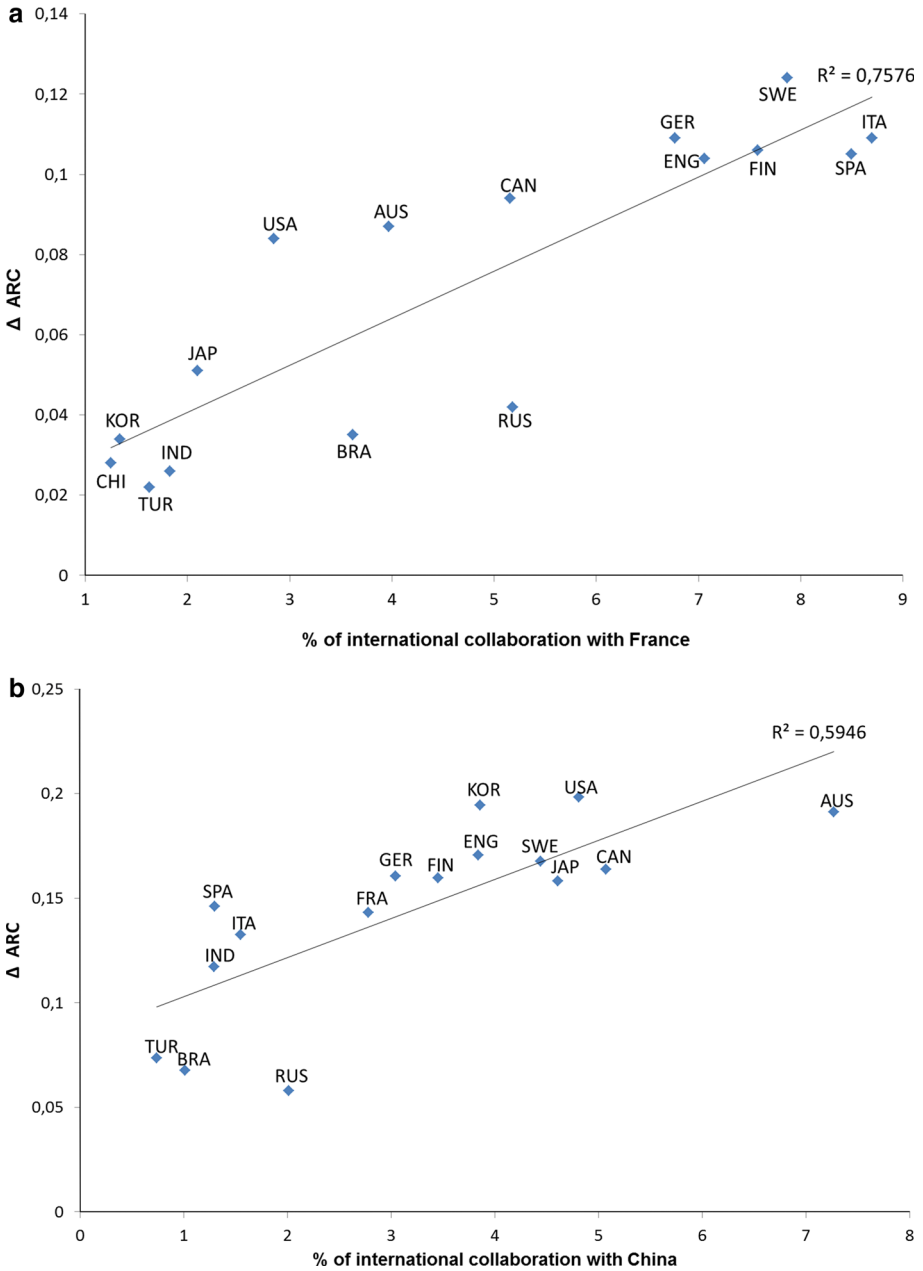


Fig. 4 **a** Relationship between Δ ARC and % of collaboration with France (2010). **b** Relationship between Δ ARC and % of collaboration with China (2010)

receives fewer than expected citations from the USA (6.2%). The trend line on Fig. 5 indicates that for each 1% of publication in collaboration with a given country, the USA addresses approximately 2% of its citations to this country. In the case of China, 4.8% of USA collaboration papers should translate in 8.6% of citations instead of 6.2%. China’s

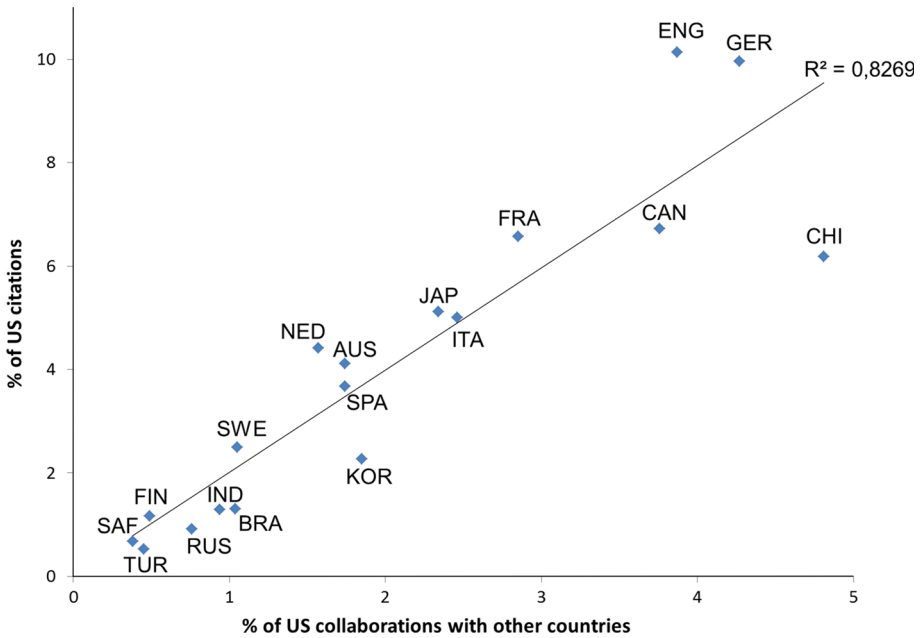


Fig. 5 Relationship between the percentage of US citations and international collaborations

position contrasts with England’s with which the USA shares “only” 3.9% of its world NSE publications. If England’s position was exactly on the trend line, it should have received 7.8% of US citations, while in fact England receives 10.1% of US citations. The difference between the two countries is explained by the fact that England’s publications have a normalized visibility factor in the USA ($VIS_{USA \rightarrow ENG}$) of 1.65 comparatively to 0.49 for China. This means that, despite the presence of a strong correlation, collaboration with the USA does not always translate in the expected number of citations. In fact, the number of citations received varies among countries. Anglo Saxon and European countries (England, Australia, Sweden, Netherlands, France), situated above the trend line, are more advantaged and benefit from an overflow of USA citation comparatively to emerging nations (China, South Korea, Brazil, Turkey, Russia) that suffer from a deficit of USA citations relatively to the intensity of their collaboration with the USA.

Discussion and Conclusion

The data presented in this paper show that the exclusion of USA citations has a significant effect on the measure of other countries’ scientific impact. In fact, we verified that excluding USA citations has a more important effect on England, Canada, Germany and France drop in scientific impact than excluding citations from their own country. This effect modifies at the global scale the ranking based on ARC between England, Canada, Germany and France. Excluding USA citations has also a more pronounced effect on these four countries’ international publications than on their national publications, which is not a surprising result considering that international collaborations have more international visibility than national publications and that part of these international collaborations are

with the USA. Of course, the drop in ARC is more pronounced when excluding USA rather than other countries' citations, because the share of USA publications in the WoS database is larger and, hence, the probability of being cited by a USA publication is higher. However, the probability of being cited by a USA publication is not evenly distributed among other countries. Some have a greater citation potential than others, as demonstrated by the differences in the calculated values of visibility factors which were higher for England and Canada than for Germany and France. This citation advantage was verified by the high correlation found between the drop in scientific impact of our selected list of 16 countries and the percentage of these countries' collaborations with the USA. It was further confirmed by the high correlation found between the percentage of USA international collaboration with other countries and the percentage of USA citations addressed to these same countries.

Considering these high correlations and the overrepresentation of USA publications in the WoS database (Archambault et al. 2006), this makes the USA citation advantage obvious to countries with which the USA collaborates the most and from which it is historically closer, that is to say Anglo Saxon countries. This advantage is strengthened by the national citation bias which was observed for each of the five analyzed countries in general and in the USA in particular. The USA national citation bias mirrors a “preference” of this country's publications to cite publications where at least one USA address appears. This means that countries that collaborate the most with the USA also benefit from this citation bias. Except for China and emerging countries which do not reap all the expected benefits of collaborating with the USA in terms of citation flows from this country, the rule generally holds for European and Anglo Saxon countries.

The major conclusion coming out of this study is that, given the strong presence of the USA in the WoS database, the comparative rankings tend, by construction, to give a citation advantage to countries having the closest relation to that country. That explains, for example, that even if France has a long history of winning Nobel Prizes and Fields Medals, it is curiously always ranked below Canada a country that rarely obtains such prizes. It is certainly difficult to believe that generally speaking, Canadian Science is superior to French science. One should thus take great care when using global indicators that do not take into account the contingent facts behind the structure of a citation database. Each result should be interpreted in light of the specific properties of the database used. For this reason, a similar exercise should be done with Scopus to better understand its internal structure in order to be able to adequately interpret the meaning of the indicators obtained from it as compared to those obtained with the WoS.

In short, and as a general principle, one should always start thinking before counting. For it is only in this manner that we will be able to realize the wish of Eugene Garfield of preventing the abuse of SCI and, more generally bibliometrics, by “devoting the necessary attention to its proper and judicious exploitation” (Garfield 1970).

References

- Aksnes, D. (2003). Characteristics of highly cited papers. *Research Evaluation*, 12(3), 159–170.
- Archambault, E., Vignola-Gagne, E., Coté, G., Larivière, V., & Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics*, 68(3), 329–342.
- Bar-Ilan, J. (2008). Informetrics at the beginning of the 21st century—A review. *Journal of Infometrics*, 2(1), 1–52.

- Basken, P. (2016). As concerns grows about using data to measure faculty, a company changes its message. *The Chronicle of Higher Education*, 11 October 2016. <http://www.chronicle.com/article/As-Concerns-Grow-About-Using/238034>.
- Bordon, M., Gomez, I., Teresea Fernandez, M., Angeles Zulueta, M., & Mendez, A. (1996). Local, domestic and international scientific collaboration in biomedical research. *Scientometrics*, 37(2), 279–295.
- Bormann, L., & Leydesdorff, L. (2013). Macro-indicators of citation impacts of six prolific countries: InCites data and the statistical significance of trends. *PLoS ONE*, 8(2), e56768.
- Collazo-Reyes, F. (2014). Growth of the number of indexed journals of Latin America and the Caribbean: The effect on the impact of each country. *Scientometrics*, 98(1), 197–209.
- Confraria, H., & Godinho, M. (2015). The impact of African science: A bibliometric analysis. *Scientometrics*, 102(2), 1241–1268.
- Davidson Frame, J., & Carpenter, M. K. (1979). International research collaboration. *Social Studies of Science*, 9(4), 481–497.
- de Beaver, B. (2004). Does collaborative research have greater epistemic authority? *Scientometrics*, 60(3), 399–408.
- European Commission. (2014). *Research and innovation performance in EU member states and associated countries. Union progress at country level*. Brussels: European Commission.
- Frenken, K., Hölzl, W., & de Vor, F. (2005). The citation impact of research collaborations: The case of European biotechnology and applied microbiology (1988–2002). *Journal of Engineering and Technology Management*, 22(1), 9–30.
- Garfield, E. (1970). Citation indexing for studying science. *Nature*, 227(5259), 669–671.
- Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics*, 51(1), 69–115.
- Glänzel, W., Debackere, K., & Meyer, M. (2008). ‘Triad’ or ‘tetrad’? On global changes in a dynamic world. *Scientometrics*, 74(1), 71–88.
- Gonzalez-Brambila, C. N., Reyes-Gonzalez, L., Veloso, F., & Perez-Angón, M. A. (2016). The scientific impact of developing nations. *PLoS ONE*, 11(3), e0151328.
- Guerrero-Bote, V. P., Olmeda-Gomez, C., & de Moya-Anegón, F. (2013). Quantifying the benefits of international scientific collaboration. *Journal of the Association for Information Science and Technology*, 64(2), 392–404.
- Gul, S., Nisa, N. T., Shah, T. A., Gupta, S., Jan, A., & Ahmad, S. (2015). Middle East: Research productivity and performance across nations. *Scientometrics*, 105(2), 1157–1166.
- Hsu, J., & Huang, D. (2011). Correlation between impact and collaboration. *Scientometrics*, 86(2), 317–324.
- Inhaber, H., & Alvo, M. (1978). World science as an input-output system. *Scientometrics*, 1(1), 43–64.
- Jurajda, S., Kozubek, S., Münich, D., & Skoda, S. (2017). Scientific publication performance in post-communist countries: Still lagging far behind. *Scientometrics*, 112(1), 315–328.
- Katz, J. S. (1994). Geographical proximity and scientific collaboration. *Scientometrics*, 31(1), 31–43.
- King, D. A. (2004). The scientific impact of nations. *Nature*, 430(6997), 311–316.
- Lancho-Barrantes, B., Guerrero-Bote, V. P., & de Moya-Anegón, F. (2013). Citation increments between collaborating countries. *Scientometrics*, 94(3), 817–831.
- Larivière, V., & Gingras, Y. (2010). The impact factor’s Matthew Effect: A natural experiment in bibliometrics. *Journal of the American Society for Information Science and Technology*, 61(2), 424–427.
- Larivière, V., Gingras, Y., & Archambault, E. (2006). Canadian collaboration networks: A comparative analysis of the natural sciences. *Social Sciences and the Humanities. Scientometrics*, 68(3), 519–533.
- 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.
- Leydesdorff, L., & Wagner, C. S. (2009). Is the United States losing ground in science? A global perspective on the world science system. *Scientometrics*, 78(1), 23–36.
- Leydesdorff, L., Wagner, C. S., & Bormann, L. (2014). The European Union, China, and the United States in the top-1% and top-10% layers of most-frequently cited publications: competition and collaborations. *Journal of Informetrics*, 8(3), 606–617.
- Leydesdorff, L., & Zhou, P. (2005). Are the contributions of China and Korea upsetting the world system of science? *Scientometrics*, 63(3), 617–630.
- Lozano, G. A., Larivière, V., & Gingras, Y. (2012). The weakening relationship between the impact factor and papers’ citations in the digital age. *Journal of the Association for Information Science and Technology*, 63(11), 2140–2145.
- Luukkonen, T., Persson, O., & Sivertsen, G. (1992). Understanding patterns of international collaboration. *Science, Technology and Human Values*, 17(1), 101–126.
- May, R. (1997). The scientific wealth of nations. *Science*, 275(5301), 793–796.

- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, *106*(1), 213–228.
- Narin, F., Stevens, K., & Whitlow, E. S. (1991). Scientific co-operation in Europe and the citation of multinationally authored papers. *Scientometrics*, *21*(3), 313–323.
- National Science Board. (2016). *Science and engineering indicators 2016*. Arlington, VA: National Science Foundation.
- Öquist, G., & Benner, M. (2015). Why are some nations more successful than others in research impact? A comparison between Denmark and Sweden. In I. Welpel et al. (Eds.), *Incentives and performance* (pp. 241–257). Springer: Governance of Research Organizations.
- Rabkin, Y., & Inhaber, H. (1979). Science on the periphery: A citation study of three less developed countries. *Scientometrics*, *1*(3), 261–274.
- Rodriguez-Navarro, A., & Narin, F. (2017). European paradox or delusion—Are European science and economy outdated? *Science and Public Policy*. <https://doi.org/10.1093/scipol/scx021>.
- Sarwar, R., & Hassan, S. (2015). A bibliometric assessment of scientific productivity and international collaboration of the Islamic World in science and technology (S&T) areas. *Scientometrics*, *105*(2), 1059–1077.
- Van Leeuwen, T. N., Moed, H. F., Tijssen, R. J., Visser, M. S., & Van Raan, A. F. (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics*, *51*(1), 335–346.
- Witze, A. (2016). Research gets increasingly international. *Nature News*, 19 January 2016. <http://www.nature.com/news/research-gets-increasingly-international-1.19198>.
- Zhou, P., & Leydesdorff, L. (2006). The emergence of China as a leading nation in science. *Research Policy*, *35*(1), 83–104.
- Zitt, M., Bassecoulard, E., & Okubo, Y. (2000). Shadows of the past in international collaboration: Collaboration profiles of the top five producers of Science. *Scientometrics*, *47*(3), 627–657.