

A comparison of the scientific performance of the U.S. and the European Union at the turn of the 21st century

Pedro Albarrán · Juan A. Crespo · Ignacio Ortuño ·
Javier Ruiz-Castillo

Received: 24 November 2009 / Published online: 20 April 2010
© Akadémiai Kiadó, Budapest, Hungary 2010

Abstract In this paper, scientific performance is identified with the impact that journal articles have through the citations they receive. In 15 disciplines, as well as in all sciences as a whole, the EU share of total publications is greater than that of the U.S. However, as soon as the citations received by these publications are taken into account the picture is completely reversed. Firstly, the EU share of total citations is still greater than the U.S. in only seven fields. Secondly, the mean citation rate in the U.S. is greater than in the EU in every one of the 22 fields studied. Thirdly, since standard indicators—such as normalized mean citation ratios—are silent about what takes place in different parts of the citation distribution, this paper compares the publication shares of the U.S. and the EU at every percentile of the world citation distribution in each field. It is found that in seven fields the initial gap between the U.S. and the EU widens as we advance towards the more cited articles, while in the remaining 15 fields—except for Agricultural Sciences—the U.S. always surpasses the EU when it counts, namely, at the upper tail of citation distributions. Finally, for all sciences as a whole the U.S. publication share becomes greater than that of the EU for the top 50% of the most highly cited articles. The data used refers to 3.6 million articles published in 1998–2002, and the more than 47 million citations they received in 1998–2007.

Keywords Research performance · Citation analysis · Scientific ranking · European paradox

Introduction

This paper compares the scientific performance of the U.S. and the EU, namely, the 15 countries forming the European Union before the 2004 accession. Like all other

P. Albarrán · I. Ortuño · J. Ruiz-Castillo (✉)
Departamento de Economía, Universidad Carlos III, Madrid, Spain
e-mail: jrc@eco.uc3m.es

J. A. Crespo
Departamento de Economía Cuantitativa, Universidad Autónoma de Madrid, Madrid, Spain

contributions to the literature referenced below, scientific performance is identified with the citation impact achieved by articles published in more than 8,000 academic or professional journals in 36 languages indexed by Thomson Scientific (TS hereafter), previously known as the Institute for Scientific Information (ISI). As far as data is concerned, the difference with other studies is that this paper uses a large sample consisting of about 3.6 million articles published in 1998–2002, as well as the approximately 47 million citations they receive in 1998–2007. The articles belong to the 20 natural sciences and the two social sciences distinguished by TS.

The design of a good science and technology policy for any area should start from an accurate diagnosis of the situation. The following two facts are well known. Firstly, since the mid-1990s the EU has had somewhat more publications than the U.S. but in terms of citations it has had a lower impact. Secondly, among highly cited papers the EU performs particularly badly. This paper illustrates these facts at the turn of the 20th century in light of both standard and novel output indicators of scientific performance. In particular, the distinctive feature of our contribution is the computation of the EU and U.S. publication shares at a large number of percentiles of the world citation distribution in 22 scientific fields.

In short, our evidence indicates that among the most influential articles, in 21 out of 22 fields the dominance of the U.S. over the EU is overwhelming. Thus, although the share of articles published in the period 1998–2002 in all sciences combined is greater in the EU than in the U.S., the latter surpasses the former when the top 50% of the most highly cited articles is considered.

The rest of the paper is organized in four sections. “[Literature review](#)” section contains a brief literature review. “[Sources and methods](#)” section presents the data and discusses our methods, while “[Empirical results](#)” section contains the empirical findings for each of the 22 scientific fields and all sciences as a whole. “[Discussion and conclusions](#)” section discusses the results and offers some concluding comments.

Literature review

It will be useful to separately review the official reports published by the European Commission and the academic literature.

Official European reports

The goal launched at the often-quoted 2000 Lisbon meeting by the Council of the EU (“*to become in 2010 the most competitive and dynamic knowledge-based economy in the world*”), would seem to reveal an urge to change the European ways in the face of two worrisome circumstances: a sizable scientific gap with the U.S. dating at least from the middle of the last century, and the awakening of several developing countries in Asia that will surely become formidable rivals to everyone in some scientific and technological fields early in the 21st century. As a matter of fact, the *Third European Report on Science and Technology Indicators*—which is the latest available in this series—states that, as far as the transition to a knowledge-based economy is concerned, “*At the beginning of this century, the EU is lagging behind the U.S. and Japan. Moreover, the observed growth rates will not allow Europe to catch up rapidly, certainly not by 2010*” (EC 2003a, p. 31).

When one searches for a diagnostic about the role of science in this rather pessimistic outlook, one encounters the notion of the “European Paradox” popularized in the *First*

European Report on Science and Technology Indicators (EC 1994), according to which Europe plays a leading world role in terms of scientific excellence but lacks the entrepreneurial capacity of the U.S. to transform it into innovation, growth, and jobs. Apparently, the problem lies—not in the EU’s scientific performance—but elsewhere. Some years later, the executive summary of the *Second Report* states that “Overall the EU’s scientific performance is excellent ... in 1995 European production overtook that of American researchers, registering almost 208,000 scientific publications compared to 203,000 for the U.S.” (EC 1997, p. 5). “The EU performs well compared with Japan and the U.S. in terms of scientific output (publications) per unit of related expenditure.” However, “there is some evidence to support the so-called European Paradox—that the EU’s healthy performance in scientific research is not being translated into strong technological and economic performance—but the picture varies across different member states and industries.” (EC 1997, p. 8).

There are, of course, other dimensions of scientific performance distinct from publication shares. As a matter of fact, the *Third Report* recognizes that “Despite decreasing publication shares, NAFTA publications tend to have high citation rates, and high relative citation impact records.”¹ (EC 2003a, p. 287). Similarly, the 2002 version of the *Key Figures* series leaves it perfectly clear that “When it comes to highly cited papers, all EU countries lag behind the U.S.” (EC 2002, p. 48).² Nevertheless, one gets the distinct impression that analysts working for the European Commission are systematically more interested in publication shares than in citation impact indicators. Very possibly, this is also why they present a rather optimistic view of the EU’s scientific performance. Thus, the concluding section of the *Third Report* emphasizes that “Although the picture across wider Europe is diverse, the EU-15 is now the best performing world region in terms of number of publications”; a fact only qualified by the statement “... it remains to be seen if the surge in European publications will lead to a similar rise in citation impact scores in the long run” (EC 2003a, p. 316). Similarly, among the key findings emphasized in the above quoted *Key Figures* document, one reads “In terms of scientific performance, the EU as a whole is doing well. With respect to highly cited publications, some EU countries show outstanding world shares” (EC 2002, p. 43). Finally, when a word of caution is provided in a brief document under the title *From ‘European Paradox’ to declining competitiveness?*, the concern with the deterioration of Europe’s leadership in scientific performance refers exclusively to scientific publications (EC 2003b, p. 1).

Academic literature

There is a scant but interesting academic literature on the scientific wealth of nations, based largely on the Science Citation Index (SCI) established by ISI since 1981 (see inter alia May 1997; Adams 1998; Katz 2000; Glänzel 2000; Glänzel et al. 2002; King 2004; Leydesdorff and Wagner 2009). Studies differ in the time period covered, the length of the citation window, the number of countries and broad fields covered, and the type of indicators used. However, certain common findings stand out: (i) Since the mid-1990s, the EU has published more than the U.S. This is particularly clear in King (2004), who illustrates the surpassing of the U.S. by the EU in the period 1997–2002 versus 1993–1997, as well as Leydesdorff and Wagner (2009) who covers the period 1995–2006. (ii) The previous fact

¹ NAFTA includes the U.S., Canada, and Mexico.

² As a matter of fact, for later reference highly cited papers as a percentage of total number of scientific publications are 1.64 and 0.25 in the U.S. and the EU, respectively.

partly explains why the gap in the total share of citations between the two areas is closing (King 2004; Leydesdorff and Wagner 2009).³ (iii) The U.S. appears as the dominant country or among the handful of countries with the highest relative citation impact indicators (May 1997; Adams 1998; Katz 2000; King 2004). (iv) The U.S. also appears among the countries with the greatest percentage of highly cited papers (King 2004; Leydesdorff and Wagner 2009). (v) The EU is very diverse. In particular, together with Switzerland outside the EU and the U.K. inside it, the excellent relative position of some small Western member countries—such as Finland, Denmark, the Netherlands, Sweden, and Belgium—should be emphasized. The performance of these small countries is highlighted in Glänzel (2000) and King (2004), as well as in the relational charting approach of Glänzel et al. (2002) and in the references in that paper to this group's previous work.⁴

However, in the academic literature we also find rather different summary views about the first axes of the so-called European Paradox. For example, King (2004) asserts "*The United States easily heads the list of nations in the volume of publications and citations and the share of top 1% cited papers, although the EU15 countries now publish more papers than the United States and are not far behind on citations*" (p. 311). Furthermore, "...comparing citations between the United States and the EU15 shows that the gap between the two has shrunk significantly since May (1997)'s analysis based on figures up to 1993. The EU now matches the United States in the physical sciences, engineering and mathematics, although still lags in the life sciences" (p. 316). Similarly, after reviewing the evidence from different sources, Shelton and Holdridge (2004) indicate: "*While the U.S. leads in most input indicators, output indicators may be more specific for determining present leadership. They show that the EU has taken the lead in important metrics and it is challenging the U.S. in others*" (Abstract, p. 353). These authors conclude "*So who is leading the world in Science and Technology: the U.S. or the EU? While no single nation rivals the U.S. for the lead, it is becoming clear that the European Union as a whole is mounting a serious challenge.*" (p. 362). On the other hand, as Leydesdorff and Wagner (2009) warn, "... the U.S. is still by far the leading nation in the world of science. The numeric lead of the EU-25 ... cannot hide the endogenous problems of the EU science system" (p. 24). While in their own review of the evidence, Dosi et al. (2006) forcefully argue, "*The general conclusion from the bibliometric data is therefore far from supporting any claim to European leadership in science. On the contrary, one observes a structural lag in top-level science vis-à-vis the US, together with (i) a few sectoral outliers in physical sciences and engineering, and (ii) a few single institutional outliers (such as Cambridge in computer science and a number of other disciplines). The first fact on which the 'Paradox' is supposedly based is simply not there. Rather a major EU challenge is how to catch up with the U.S. in terms of scientific excellence.*" (p. 1455).

Thus, we believe there is room for a contribution that clarifies the extent of the gap between the scientific performance of the U.S. and the EU. After briefly reviewing the standard methods used in the official and academic literature, this paper looks at a large data set from an innovative perspective to compare the relative position of the U.S. and the EU at the turn of the century in 22 natural and social sciences.

³ These two facts are corroborated in Fig. 2 of the previous version of this paper in Albarrán et al. (2009).

⁴ On the other hand, Katz (2000) adjusts relative citation impact indicators to take into account a strong, non-linear relationship between the number of citations a collection of papers receives and the collection size. As a consequence, there is a dramatic reversal of positions in many sub-fields between the U.S. and some European and non-European countries (see EC 2003a, pp. 443–444 for a large reversal between large and small countries). However, a discussion of Katz's approach is beyond the scope of this paper.

Sources and methods

Description of the sample

We start from the set of all papers published in the periodical literature indexed by TS in 1998–2007. TS-indexed papers include research articles, reviews, proceedings papers and research notes. In this paper, only research articles, or simply articles, are studied. After the elimination of observations with missing information, the initial dataset consists of 8,153,092 articles, or approximately 95% of the number of items in the original database. This dataset consists of articles published in a certain year and the citations they receive from that year until 2007, that is, articles published in 1998 and their citations during the 10-year period 1998–2007, and so on up to articles published in 2007 and their citations during that same year.

A number of scholars have highlighted fundamental differences between the publishing and citation practices of scholars in the natural sciences and those in the social sciences and the humanities (see inter alia the references in Archambault et al. 2006). On the other hand, it has often been argued that, among the TS databases, the *Social Sciences Citation Index* (SSCI) and the *Arts and Humanities Citation Index* (AHCI) do not adequately cover journal languages other than English.⁵ These features justify the usual exclusive concentration in the natural sciences. In addition, the Arts and Humanities citation distribution in our dataset presented very different characteristics from the remaining TS fields (for example, 83% of all articles received no citations at all). However, this was not the case with two social sciences: Economics and Business, and the General Social Sciences field distinguished by TS.⁶ Therefore, we decided to include these social sciences together with the 20 fields in the natural sciences distinguished by TS. The final dataset consists of 7,988,571 articles and 65,042,734 citations.

The problem of language of scientific publications does not end here. There is a general agreement that the *Science Citation Index* (SCI) suffers from an English language bias (van Leeuwen et al. 2001). However, taking into account that English can be considered the international language of science, in this paper we have followed the usual practice of using the TS data under the reasonable assumption that “*the international journal publications in these databases provide a satisfactory representation of internationally accepted ('mainstream') research, especially high-quality 'laboratory based' basic research in the natural sciences, medical sciences, and life sciences conducted in the advanced industrialized nations*” (EC 2003a, p. 439).

There are two reasons why we need a large sample of articles and the citations they receive. Firstly, this paper aims to obtain empirical conclusions at the level of 22 TS fields, but thirteen fields might be too small if we were to take only articles published in a single year. Secondly, we want to establish some stylized facts about the relative scientific performance of the U.S. and the EU at the turn of the 21st century, when the Lisbon declaration by the European Council took place. Consequently, the remaining part of the paper essentially focuses on the sample of 3,654,675 articles published in 1998–2002 and

⁵ Archambault et al. (2006) have recently established that there is a 20–25% overrepresentation of English-language journals in TS's databases compared to the list of journals in Ulrich's International Periodicals Directory.

⁶ Albarrán and Ruiz-Castillo (2009) contains a discussion of the characteristics shared by these two social sciences and the remaining broad scientific fields.

the 47,239,360 citations they receive in 1998–2007, that is, the maximum volume of citation existing in our database.

Assignment of articles to geographical areas

Articles are assigned to geographical areas according to the institutional affiliation of their authors as recorded in the TS database on the basis of what had been indicated in the by-line of the publications. The assignment of internationally co-authored papers among areas is problematic (see *inter alia* Anderson et al. 1988). From a U.S. geopolitical point of view, for example, we want to give as much weight to an article written in a U.S. research center as we give to another co-authored by researchers from a U.S. and a European university. Thus, as in all studies reviewed in “Literature review” section, in this paper in every internationally co-authored article a whole count is credited to each contributing area. Only domestic articles, or articles exclusively authored by one or more scientists affiliated to research centers either in the U.S. or the EU alone, are counted once.⁷

Note that, in the presence of geographical areas consisting of several countries, this is the best we can do with the available information. Alternatively, articles in the EU, for example, could be assigned in two steps: first to individual member countries, and then to the European aggregate. Of course, this procedure would artificially blow up the European share to the extent of intra-European cooperation among member countries. We estimate that the upward bias in publications starts in 1998 at 15% and increases with time until 2007 (see Albarrán et al. 2009).⁸

Table 1 presents the percentage distribution by field of the extended number of articles assigned to the U.S. and the EU. The 20 fields in the natural sciences are organized in three large aggregates: Life Sciences, Physical Sciences, and Other Natural Sciences, including the Multidisciplinary field that, as Tijssen and van Leeuwen (2003) explain, includes the prestigious general journals with broad multidisciplinary scopes, such as *Nature*, and *Science*. Since the two areas belong to the so-called “Western model” (see *inter alia* Glänzel 2000, 2001), the two distributions are rather close to each other. Relative to a situation in which the two areas were to be heavily specialized in different fields, this correlation makes any *a priori* differentiation among fields unnecessary.

Methods

The standard output indicators used in the literature can be briefly reviewed in two steps.⁹ In the first place, a natural performance indicator is the share of publications during a given

⁷ It should be noted that when the 1998–2002 dataset is partitioned into the U.S., the EU and a third geographical area consisting of the rest of the world, the total number of articles in such extended count is 13.6% more than the standard count in which all articles are counted once. Similarly, the total number of citations in the extended sample is 20.2% greater than the one in the standard dataset. For further details, see Albarrán et al. (2009).

⁸ This is an important measurement issue that might be affecting the important contribution by King 2004, whose Fig. 1 (p. 311) states: “The EU15 total contains some duplication because of papers jointly authored between countries in the EU group.”

⁹ For the simultaneous measure of outputs and inputs to the scientific and innovation process, as well as a discussion of productivity indicators, see May (1997, 1998), EC (2003a), King (2004), and Shelton and Holdridge (2004). The latter also includes a review of qualitative methods for the measurement of science and technology consisting of studies of the international stature of research centers in the U.S. and the EU conducted by experts in the corresponding disciplines. For a general discussion of the evolution and

Table 1 Number of extended articles, and extended distributions by field in the U.S. and the EU 1998–2002

Fields	Number of extended articles			
	EU	%	U.S.	%
Life Sciences	583,390	43.5	565,370	47.0
1. Clinical Medicine	31,086	23.6	279,776	23.3
2. Biology & Genetics	83,941	6.3	82,560	6.9
3. Neuroscience & Behavior	45,361	3.4	47,414	3.9
4. Molecular Biology	39,716	2.9	45,541	3.8
5. Psychiatry & Psychology	27,738	2.1	49,490	4.1
6. Pharmacology & Toxicology	22,676	1.7	18,479	1.5
7. Microbiology	25,545	1.9	19,967	1.7
8. Immunology	21,327	1.6	22,143	1.8
Physical Sciences	389,102	29.0	262,488	21.8
9. Chemistry	155,842	11.6	93,631	7.8
10. Physics	143,442	10.7	92,397	7.7
11. Computer Science	27,460	2.0	25,247	2.1
12. Mathematics	38,059	2.8	29,110	2.4
13. Space Science	24,299	1.8	22,103	1.8
Other Natural Sciences	316,166	23.5	267,987	22.3
14. Engineering	91,836	6.8	85,565	7.1
15. Plant & Animal Science	73,771	5.5	64,330	5.3
16. Materials Science	52,647	3.9	31,211	2.6
17. Geosciences	37,550	2.8	33,350	2.8
18. Environment & Ecology	31,684	2.4	32,118	2.7
19. Agricultural Sciences	24,824	1.8	16,557	1.4
20. Multidisciplinary	3,854	0.3	4,856	0.4
Social Sciences	52,702	3.9	106,569	8.9
21. Social Sciences, General	35,145	2.6	78,916	6.6
22. Economics & Business	17,557	1.3	27,653	2.3
All Sciences	1,341,360	100.0	1,202,414	100.0

time period. When there is information on the citations received by these publications, two other indicators are typically added: the share of total citations, and some measure of the citation impact of the average paper.¹⁰ In the second place, the problem with all these indicators is that two distributions that share the same mean might be very differently shaped away from the mean. In our case, it is well known that citation distributions are highly skewed: according to Albarrán and Ruiz-Castillo (2009), for example, about 70% of

Footnote 9 continued

shortcomings of science and technology indicators and their use in national policy, see Grupp and Mogege (2004).

¹⁰ There are two types of average-based measures: the impact measures rebased against the world baseline, used *inter alia* in May (1997), Adams (1998), King (2004), EC (2003a), and Shelton and Holdridge (2004), and the relational charts in Glänzel et al. (2002) that use information—unavailable in our database—about the journals where each country's articles are published.

all articles published in 1998–2002 with a fixed five-year citation window receive fewer than 30% of all citations while fewer than 10% of them account for more than 40% of all citations. This is surely the reason why the authors in the Leiden group have always completed the average-based indicators used to monitor research groups with the percentage of uncited papers (Moed and van Raan 1988; Moed et al. 1985, 1995).¹¹ More recently, they also include the percentage contribution in the top 5% of most highly cited papers (van Leeuwen et al. 2003). In the same vein, recent contributions to the literature on international comparisons of scientific performance—such as King (2004)—include the country's share in the top 1% of the most highly cited papers as an output indicator, information made readily available by TS since 2001 in their *Essential Science Indicators* (<http://www.isihighlycited.com>).¹²

The distinctive feature of this paper is the computation of the EU and U.S. publication shares at a large number of percentiles of the world citation distribution. The articles published in any given scientific field over a given time period are ordered in increasing number of citations. Then the shares of articles attributed to both geographical areas are computed at every percentile p in the unit interval. When $p = 0.1$, for example, the shares refer to the set of articles after discarding the 10% least cited, or what is the same, to the 90% of the most highly cited articles. For a given geographical area, the graph of the publication shares as p increases from 0.1 to 0.2, 0.3, etc., indicates its relative performance as the publications' impact measured by the number of citations increases. The comparisons of such graphs for two geographical areas provide an eloquent picture of their relative situation at many points of the citation distribution, and not only when $p = 0$ and $p = 0.01$ as is the case in recent contributions to this literature.¹³

Empirical results

Standard output indicators

Table 2 presents the following standard output indicators of scientific performance for the U.S. and the EU in every field: (i) the share of articles published in 1998–2002, (ii) the share of total citations received by these articles in 1998–2007, and (iii) the normalized mean citation rate per article.¹⁴ Scientific fields are classified in two groups: Group A (GA hereafter) includes fields where the share of articles in the U.S. is greater than the share in the EU, while Group B (GB hereafter) includes fields where the opposite is the case.

As can be observed in the last row and columns 1 and 2 in Table 2, the share of articles in all fields combined in 1998–2002 is greater in the EU than in the U.S. However, this hides differences across fields that it is important to highlight. Among GA fields there are twice as many Social Science articles in the U.S. as there are in the EU. Taking into

¹¹ The Leiden group also constructs their average-based indicators counting with information about the journals where each country's articles are published. This allows them to compare the research groups' observed mean citation with the expected behavior of the set of journals where the group is known to publish. The ratio of such expected behavior to the behavior of the journals in the entire field constitutes another interesting indicator in this case.

¹² See also Batty (2003) for a study of the pattern of spatial concentration by the highly cited scientists.

¹³ The same idea can be found in the study of domestic versus internationally co-authored papers in Glänzel (2000, 2001).

¹⁴ Since the total number of extended articles is greater than the actual number, the sum of the shares in (i) and (ii) over the partition of the world into geographical areas would add up to more than one.

Table 2 Standard output indicators of scientific performance in the U.S. and in the EU 1998–2002

Scientific Fields	Share of articles		Share of total citations		Normalized average citation	
	U.S. (1)	EU (2)	U.S. (3)	EU (4)	U.S. (5) = (3)/(1)	EU (6) = (4)/(2)
Group A						
1. Social Sciences, General	56.8	25.3	64.9	24.0	1.14	0.95
2. Economics & Business	53.5	34.0	70.8	28.4	1.32	0.84
3. Psychiatry & Psychology	54.6	30.6	63.4	29.2	1.16	0.95
4. Neuroscience & Behavior	41.2	39.4	53.3	38.7	1.29	0.98
5. Molecular Biology and Genetics	45.0	39.2	59.3	39.4	1.32	1.00
6. Immunology	43.4	41.8	53.3	40.3	1.23	0.97
7. Multidisciplinary	26.0	20.7	53.9	28.4	2.07	1.38
Group B						
8. Agricultural Sciences	24.7	37.0	31.5	44.1	1.28	1.19
9. Materials Science	19.2	32.5	28.8	35.3	1.49	1.09
10. Plant & Animal Science	29.9	34.3	37.0	40.0	1.24	1.17
11. Chemistry	20.8	34.6	32.3	38.8	1.55	1.12
12. Physics	24.7	38.4	38.7	44.5	1.56	1.16
13. Mathematics	30.5	39.8	41.4	44.0	1.36	1.10
14. Microbiology	33.1	42.4	44.3	44.8	1.34	1.06
15. Pharmacology & Toxicology	29.3	35.9	38.2	38.3	1.30	1.07
16. Clinical Medicine	35.9	40.7	48.9	40.6	1.36	1.00
17. Space Science	44.9	49.4	63.7	53.3	1.42	1.08
18. Computer Science	35.1	38.3	51.6	36.2	1.47	0.95
19. Geosciences	34.5	38.8	47.8	43.2	1.39	1.11
20. Engineering	29.7	31.9	37.7	35.4	1.27	1.11
21. Biology and Biochemistry	36.3	37.0	49.6	37.4	1.36	1.01
22. Environment & Ecology	36.3	35.8	43.3	38.6	1.20	1.08
Social Sciences	55.9	27.6	66.9	25.5	1.20	0.92
Physical Sciences	25.2	37.4	37.9	42.0	1.50	1.12
Other Natural Sciences	28.6	33.8	38.2	38.7	1.34	1.15
Life Sciences	38.0	39.2	51.0	39.3	1.34	1.00
All Sciences	32.9	36.7	46.3	39.5	1.41	1.08

account what was said about these disciplines in “[Description of the sample](#)” section, this large quantitative superiority of the U.S. over the EU should come as no surprise. Something similar can be said about Psychology and Psychiatry, and even Neuroscience and Behavioral Sciences. Among the GB fields, the difference between the EU and the U.S. shares declines until they became practically equal.

The key fact is that as soon as we turn from the sheer production of scientific articles toward the impact they have in terms of total citations received, the relative situation of the two geographical areas is dramatically reversed: for all sciences combined, the share of total citations of U.S. articles in our 1998–2002 dataset is almost seven percentage points greater than the EU share (see the last row and columns 3 and 4 in [Table 2](#)). This is partly the result of the clear U.S. superiority in all GA fields already discussed. In the last seven

GB fields, where the EU articles share is quite close or equal to the U.S. one, the total citation share is greater in the U.S. Only in the remaining eight, where there is considerably more EU than U.S. publications, the EU citation share is equal to or greater than that of the U.S. From another point of view, the total citation share is much greater in the U.S. for the Social Sciences and the Life Sciences, about four percentage points greater in the EU for the Physical Sciences, and equal for the Other Natural Sciences.

Finally, when we focus on the normalized mean citation rate (MCR hereafter) in columns 5 and 6 in Table 2, the comparison becomes completely one-sided: for articles published during 1998–2002 the U.S. dominates the EU in every one of the 22 scientific fields. Three facts should be noted. Firstly, except for the Multidisciplinary field, the MCR in all GA fields in the EU is just equal to or smaller than that of the world as a whole. Secondly, within GB, the MCR for the EU is considerably greater than that of the world in only eight fields. Thirdly, the EU's performance is also particularly dismal, just below or above the world standard, in three other GB fields. Looking at GB as a whole, it can be concluded that the EU publishes too many poorly cited articles. As a result, the MCR for all sciences combined in the EU is only 8% above the world one and 33 percentage points below the U.S. one.

The total dominance by the U.S. in MCRs is lost in the influential paper by King (2004). As indicated in “Literature review” section, this author states that “*the EU now matches the United States in the physical sciences, engineering and mathematics, although still lags in the life sciences*”. But this statement refers to the share of total citations (Fig. 4, p. 315), a fact essentially confirmed in columns (3) and (4) of our Table 2. The reason is that, as can be seen in columns (1) and (2) in Table 2, the number of publications in these fields is greater in the EU. However, once the number of articles is also taken into account, all MCRs become greater in the U.S.

Publication shares at different percentiles of the citation distribution

Recall that articles published in 1998–2002 receive citations over a different time period. Therefore, in order not to discriminate in favor of earlier published articles, we first partition the articles published in each of these 5 years in percentiles, and then construct each percentile for the entire period by adding up the corresponding articles published in each year.¹⁵ Finally, in each percentile, we compute the share of (extended) articles with at least one author working in a research institution in the U.S., the EU or the RW with respect to the total actual number of articles in each percentile.¹⁶ It is important to indicate that, in practice, only the following 21 percentiles are actually computed in each case: 0, 10, 15, ..., 90, 95, 98, and 99; the graphical representation in Figs. 1, 2, 3, 4 are then achieved by linear interpolation.

Figure 1 includes the seven GA fields. As can be observed, in every case the initial gap between the U.S. and the EU shares widens as p increases. Figure 2 includes eight GB fields in which the initial publication share is greater in the EU than in the U.S. The distinctive feature here is that the U.S. surpasses the EU rather early in six cases (Biology and Biochemistry at $p = 0.10$; Space Science at $p = 0.25$; Environment and Ecology at $p = 0.35$; Clinical Medicine at $p = 0.36$; Computer Science at $p = 0.46$); or at intermediate points (Geosciences at $p = 0.55$; Microbiology at $p = 0.70$, and Neuroscience and

¹⁵ This is also the method followed in the construction of the top 1% of the most highly cited articles in the Web of Science's *Essential Science Indicators*.

¹⁶ Like before, the sum of such shares at every percentile will not add up to one.

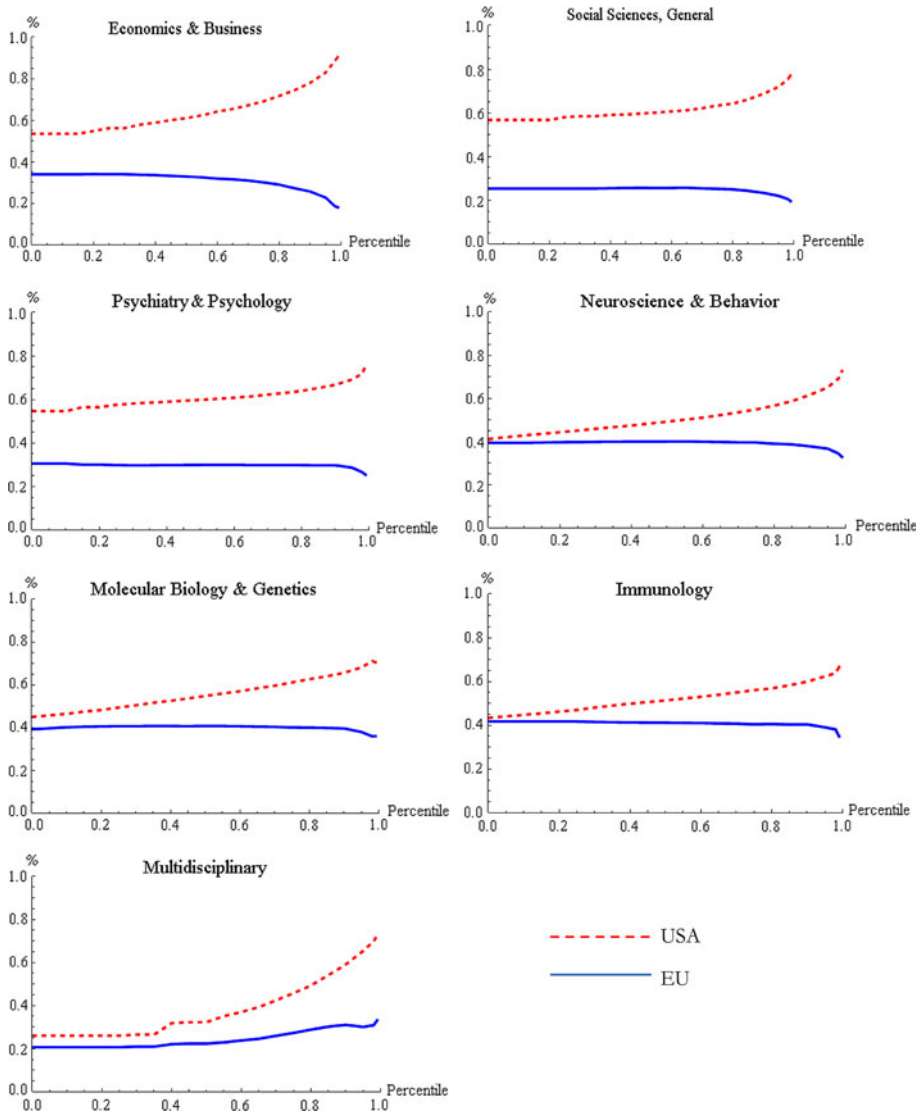


Fig. 1 Very clear U.S. superiority over the EU

Behavior at $p = 0.71$). Figure 3 includes six GB fields in which the U.S. surpasses the EU rather late (Chemistry at $p = 0.90$; Mathematics at $p = 0.92$; Engineering at $p = 0.94$; Physics and Material Sciences at $p = 0.95$, and Plant and Animal Sciences at $p = 0.96$). Finally, Fig. 4 contains the Agricultural Sciences, the only case in which the EU dominates the U.S. at every percentile, as well as all sciences combined in which the U.S. surpasses the EU at about $p = 0.50$.

Two final points are in order. Firstly, the average results in terms of MCRs and the relative performance illustrated in Figs. 1, 2, 3, 4 are consistent with each other. Secondly,

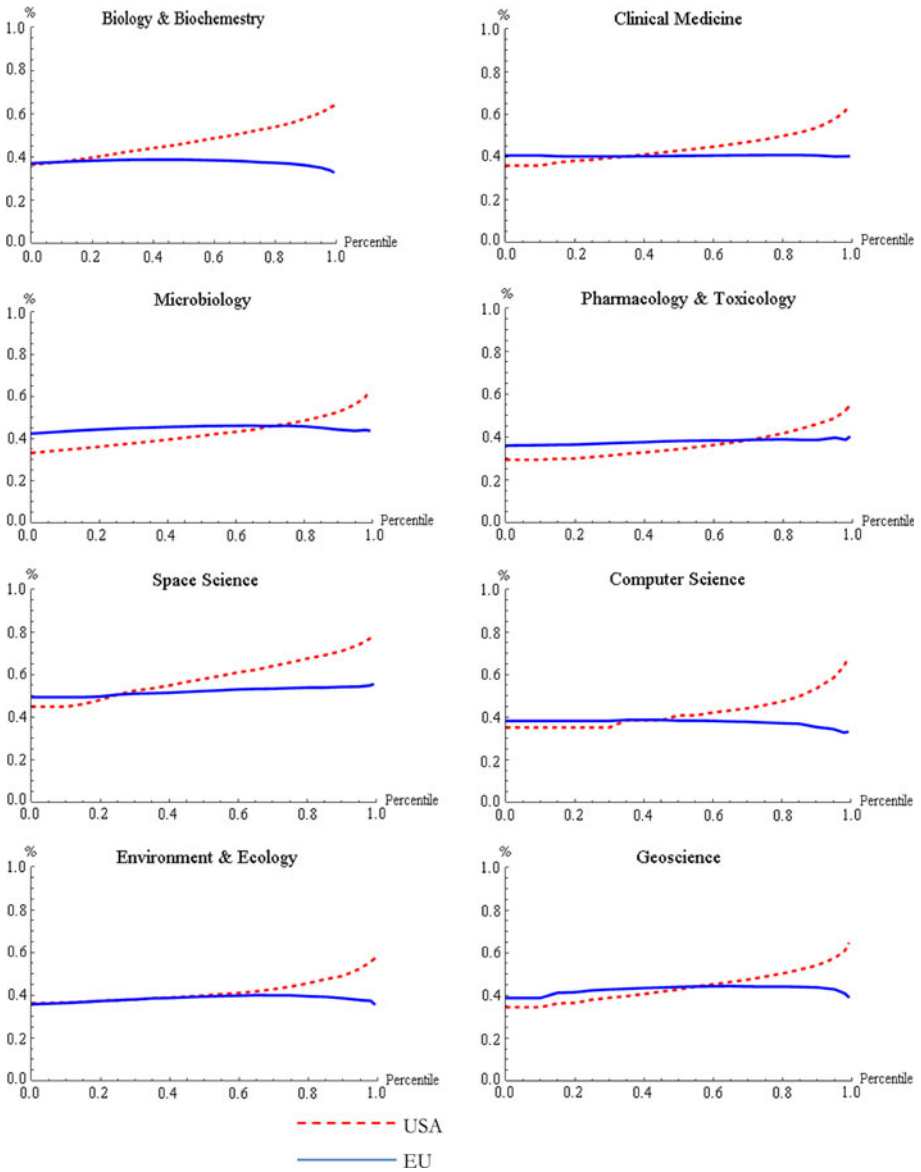


Fig. 2 Relatively clear U.S. superiority over the EU

the U.S. curves tend to have a positive slope and, when the upper tail is reached at $p = 0.90$, they all clearly rise without exception. However, in about 10 fields the EU share remains relatively flat or slightly increases, while in the remaining 12 decreases at that crucial stage. This is a point that can hardly have been ascertained with the sole help of average-based indicators.

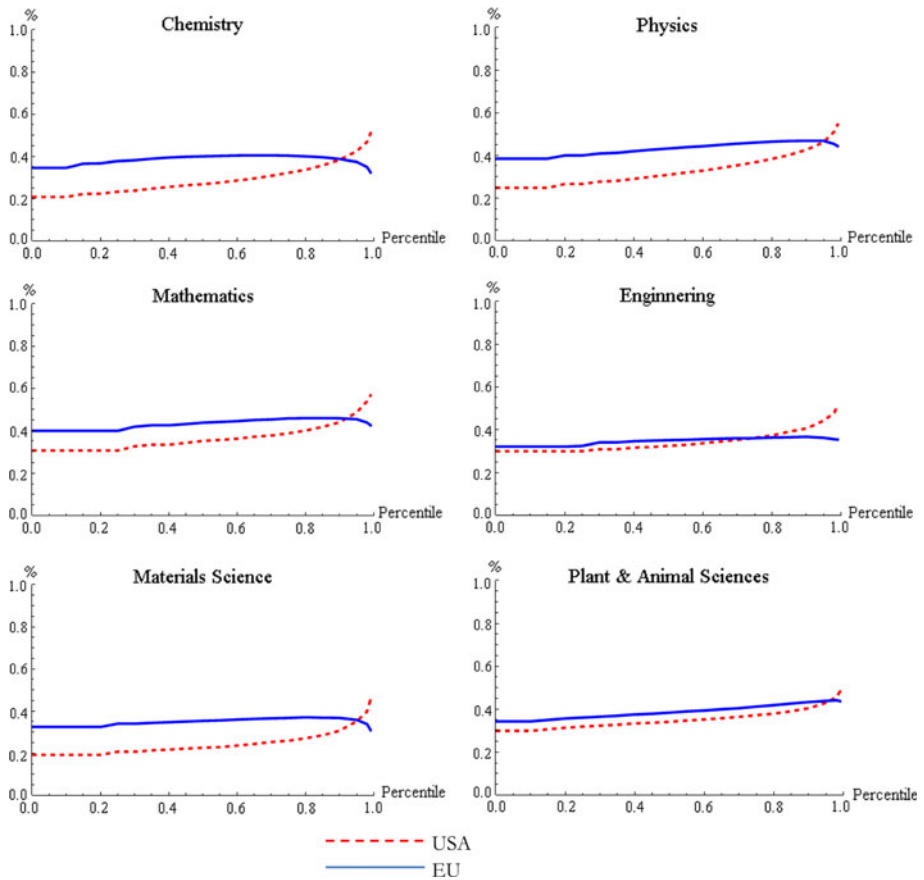


Fig. 3 Slight U.S. superiority over the EU

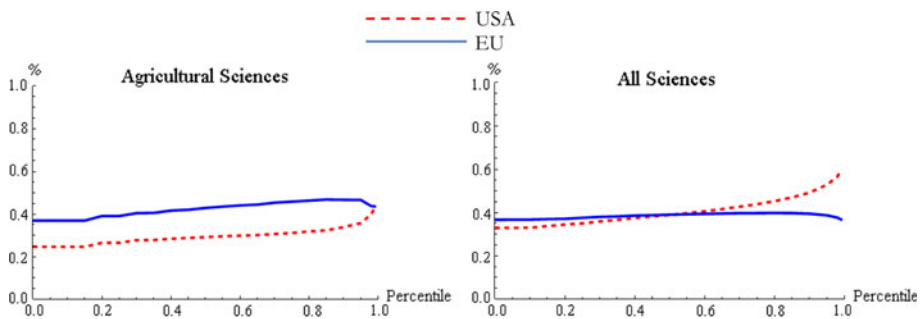


Fig. 4 The only field with EU superiority over the U.S., and the case for all sciences combined

Discussion and conclusions

As documented in “Literature review” section, different people and institutions held rather different views about the relative scientific performance of the EU and the U.S. at the turn

of the 21st century. As a contribution to the settlement of this issue, together with standard output indicators such as the MCR, this paper has compared the U.S. and EU publication shares at every percentile of the world citation distribution. The data used is a sample of 3.6 million articles published in 1998–2002 in 22 scientific fields and the more than 47 million citations they received in 1998–2007.

Before continuing any further, it is important to comment on some of the limitations of this paper. Firstly, the number of citations received by an article can be used as a proxy of the scientific visibility and influence it has achieved, which in turn partially reflects the scientific utility of the paper. However, as has often been stressed in the literature, citation data provide only a limited and incomplete view of research quality (see *inter alia* the recent report by Adler et al. 2008). As eloquently explained in Moed et al. (1985), it is virtually impossible to operationalize the general notion of “basic quality” that refers to the cognitive, methodological, and esthetic aspects of quality which can only be assessed by criteria intrinsic to scientific research within the peer review system. In this view, scientific quality in a broader sense “includes basic quality as well as the extent to which researchers successfully perform ‘public relations’ activities” (p. 135). Citation-based, impact indicators are then quality indicators in this—nevertheless important—broad sense. Secondly, it should be remembered that—as has been already stressed in “[Description of the sample](#)” section—the set of TS journals is not always a good reflection of all worldwide scientific publication output and research activity. This might very well influence the results obtained in two different directions. Some might argue that as far as the Social Sciences, and perhaps also as far as Psychology and Psychiatry and the Behavioral Sciences are concerned, the TS database favors the U.S. versus the EU.¹⁷ However, as far as the life and other natural sciences are concerned, the remarkable findings by van Leeuwen et al. (2001) indicate that countries such as Germany, France, and Switzerland may have a low citation impact score because they have a decreasing though still significant number of publications in non-English journals covered by the SCI that have a considerably lower impact than the English-language journals. Thus, when the publications in these non-English journals, but not their citations to articles in English, are removed from the publication output, the impact score of these countries show increases above at least 10%! Therefore, in comparing the citation performance of English and non-English countries it might be desirable to eliminate non-English journals altogether—a possibility beyond the scope of this paper because our dataset lacks information on journals. In any case, it might be worthwhile to redo the analysis with other databases such as Scopus.

The facts of the matter covered in this paper can be summarized in three points. Firstly, the EU share of total publications in all sciences in 1998–2002 is about 4% greater than that of the U.S. Secondly, as soon as these articles’ impact, measured by the citations they receive, is taken into account, the overall picture is reversed: the EU MCR for all sciences combined is only 8% above the world rate, but 33 percentage points below that of the U.S. Moreover, the U.S. publication share becomes greater than the EU’s for the top 50% of the most highly cited articles. Thirdly, there are of course differences across fields. In particular, the EU performs well above the world average in eight fields. However, the European MCR is considerably smaller than the U.S. one in all of these favorable cases, and the EU publication share in all these fields is surpassed by the U.S. share for all percentiles beyond the top 45% or the top 4% of the most highly cited articles, depending

¹⁷ As economists and/or members of Economics Departments, we believe that members of the European Economic Association and many other colleagues in Economics accept the information in the SCI and the SSCI as valid in our field.

on the case. On the other hand, relative to the world and the U.S. the EU performance is particularly poor in the following cases: the Social Sciences—nor surprisingly, given the nature of the data and of the fields themselves—Computer Science, and all life sciences independently of whether the total publication share is greater in the U.S. or the EU.

The overall conclusion is inescapable. Independently of sectoral details just discussed, according to our large 1998–2002 dataset acquired from TS one must fully side with Leydesdorff and Wagner (2009) and Dosi et al.'s (2006) diagnosis when they argue that “*The numeric lead of the EU-25 ... cannot hide the endogenous problems of the EU science system*”, and “*one observes a structural lag in top-level science vis-à-vis the US*”.

This is not the place for the formulation of policy recommendations. Before doing that, the present analysis should be extended in rather obvious directions. Firstly, towards specific sub-fields, countries within the EU, and even individual research centers. Secondly, it would be important to separately analyze domestic and internationally co-authored articles. In the European case, the latter should differentiate between intra-European cooperation and cooperation with the U.S. and the rest of the world. Thirdly, we also need input indicators at the lowest aggregation level of the sort referred to in note 9. Fourthly, this paper has looked at the entire citation distribution in each field. However, there is no doubt that the most relevant basic and applied research is generally found at the very top of citation distributions. This leads us to stress the following two points. On one hand, robust measures of scientific excellence, such as the *h*-index and its many variants may be particularly appropriate. On the other hand, the differential behavior of the U.S. and the EU scientific communities at the upper tail of the citation distribution in all fields calls for some systematic explanation from which to extract policy implications.

Acknowledgements The authors acknowledge financial support from the Spanish MEC through grants SEJ2007-63098, SEJ2006-05710, SEJ2007-67135, and SEJ2007-67436. The database of Thomson Scientific (formerly Thomson-ISI; Institute for Scientific Information) has been acquired with funds from Santander Universities Global Division of *Banco Santander*. This paper is part of the SCIFI-GLOW Collaborative Project supported by the European Commission's Seventh Research Framework Programme, Contract number SSH7-CT-2008-217436. Suggestions by a referee helped to improve a previous version of the paper.

References

- Adams, J. (1998). Benchmarking international research. *Nature*, 396, 615–618.
- Adler, R., Erwing, J., & Taylor, P. (2008). *Citation Statistics*, a report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS).
- Albarrán, P., & Ruiz-Castillo, J. (2009). References made and citations received by scientific articles. Working Paper 09-81, Economics Series 45, Universidad Carlos III.
- Albarrán, P., Crespo, J., Ortuño, I., & Ruiz-Castillo, J. (2009). A Comparison of the Scientific Performance of the U.S. and Europe at the Turn of the XX Century. Working Paper 09-55, Economics Series 34, Universidad Carlos III (<http://www.eco.uc3m.es/personal/cv/jrc.html>).
- Anderson, J., Collins, P., Irvine, J., Isard, P., Martin, B., Narin, F., et al. (1988). On-line approaches to measuring national scientific output: A cautionary tale. *Science and Public Policy*, 15, 153–161.
- Archambault, E., Vignola-Gagne, E., Côté, G., Larivière, V., & Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics*, 68, 329–342.
- Batty, M. (2003). The geography of scientific citation. *Environment and Planning A*, 35, 761–770.
- Dosi, G., Llerena, P., & Sylos Labini, M. (2006). Science-technology-industry links and the ‘European Paradox’: Some notes on the dynamics of scientific and technological research in Europe. *Research Policy*, 35, 1450–1464.

- EC. (1994). *First European report on science and technology indicators*. Luxembourg: Directorate-General XII, Science, Research, and Development, Office for Official Publications of the European Community.
- EC. (1997). *Second European report on science and technology indicators*. Luxembourg: Directorate-General XII, Science, Research, and Development, Office for Official Publications of the European Community.
- EC. (2002). *Key Figures. Towards a European research area. Science, technology, and innovation*. Luxembourg: Research Directorate General, Office for Official Publications of the European Community.
- EC. (2003a). *Third European report on science and technology indicators*. Directorate-General for Research. Luxembourg: Office for Official Publications of the European Community. <http://www.cordis.lu/rd2002/indicators/home.html>.
- EC. (2003b). From 'European Paradox' to declining competitiveness? Snapshots, 4. In *Key Figures 2003/2004*, Directorate-General for Research. Luxembourg: Office for Official Publications of the European Community, <http://cordis.europa.eu/indicators/publications.htm>.
- Glänzel, W. (2000). Science in Scandinavia: A bibliometric approach. *Scientometrics*, 48, 121–150.
- Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics*, 51, 69–115.
- Glänzel, W., Schubert, A., & Braun, T. (2002). A relational charting approach to the world of basic research in twelve science fields at the end of the second millennium. *Scientometrics*, 55, 335–348.
- Grupp, H., & Mogege, M. E. (2004). Indicators for National Science and Technology Policy: How robust are composite indicators? *Research Policy*, 33, 1373–1384.
- Katz, J. S. (2000). Scale-independent indicators and research evaluation. *Science and Public Policy*, 27, 23–36.
- King, D. (2004). The scientific impact of nations. *Nature*, 430, 311–316.
- Leydesdorff, L., & Wagner, C. (2009). Is the United States losing ground in science? A global perspective on the world science system. *Scientometrics*, 78, 23–36.
- May, R. (1997). The scientific wealth of nations. *Science*, 275, 793–796.
- May, R. (1998). The scientific investments of nations. *Science*, 281, 879–880.
- Moed, H. F., & van Raan A. F. J. (1988). Indicators of research performance. In A. F. J. van Raan (ed.), *Handbook of quantitative studies of science and technology* (pp. 177–192). North Holland.
- Moed, H. F., Burger, W. J., Frankfort, J. G., & van Raan, A. F. J. (1985). The use of bibliometric data for the measurement of university research performance. *Research Policy*, 14, 131–149.
- Moed, H. F., De Bruin, R. E., & van Leeuwen, Th. N. (1995). New bibliometrics tools for the assessment of national research performance: Database description, overview of indicators, and first applications. *Scientometrics*, 133, 381–422.
- Shelton, R., & Holdridge, G. (2004). The US-EU race for leadership of science and technology: Qualitative and quantitative indicators. *Scientometrics*, 60, 353–363.
- Tijssen, R., & van Leeuwen, T. (2003). *Bibliometric analysis of world science*. Extended technical annex to chapter V of EC (2003a).
- Van Leeuwen, T., Moed, H., Tijssen, R., Visser, M., & van Raan, A. (2001). Language biases in the coverage of the science citation index and its consequences for international comparisons of national research performance. *Scientometrics*, 51, 335–346.
- Van Leeuwen, T., Visser, M., Moed, H., Nederhof, T., & van Raan, A. (2003). The holy grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics*, 57, 257–280.