Journal of Informetrics 8 (2014) 606–617

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

The European Union, China, and the United States in the top-1% and top-10% layers of most-frequently cited publications: Competition and collaborations

Loet Leydesdorff[a,*], Caroline S. Wagner[b], Lutz Bornmann[c]

[a] Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Kloveniersburgwal 48, 1012 CX Amsterdam, The Netherlands
[b] Battelle Center for Science & Technology Policy, John Glenn School of Public Affairs, Ohio State University, Columbus, OH, USA
[c] Division for Science and Innovation Studies, Administrative Headquarters of the Max Planck Society, Hofgartenstr. 8, 80539 Munich, Germany

ARTICLE INFO

Article history:
Received 1 April 2014
Received in revised form 8 May 2014
Accepted 14 May 2014

Keywords:
World share
Citation analysis
Excellence
Europe
USA
China

ABSTRACT

The percentages of shares of world publications of the European Union and its member states, China, and the United States have been represented differently as a result of using different databases. An analytical variant of the Web-of-Science (of Thomson Reuters) enables us to study the dynamics in the world publication system in terms of the field-normalized top-1% and top-10% most-frequently cited publications. Comparing the EU28, USA, and China at the global level shows a top-level dynamic that is different from the analysis in terms of shares of publications: the United States remains far more productive in the top-1% of all papers; China drops out of the competition for elite status; and the EU28 increased its share among the top-cited papers from 2000 to 2010. Some of the EU28 member states overtook the United States during this decade; but a clear divide remains between EU15 (Western Europe) and the Accession Countries. Network analysis shows that China was embedded in this top-layer of internationally co-authored publications. These publications often involve more than a single European nation.

1. Introduction

In 2011, the Royal Society—the national science academy of the UK—using data from the Scopus database (an Elsevier product) issued a report showing that China was on a trend line to overtake the USA in terms of numbers of publications by 2013 (Clarke & Plume, 2011; Plume, 2011). Along the same lines, Hill, Rapoport, Lehming, and Bell (2007) and Wagner (2011) showed that several European nations had increased their overall citation shares, and six European countries had overtaken the USA in terms of relative citation rates. In this study, we explore these trends further by examining patterns among the most-highly cited papers, expecting to find that country shares among the most elite papers reflect specific historical patterns (Bornmann, de Moya Anegón, & Leydesdorff, 2010). Moreover, we expect to find that the most elite scientists are highly networked internationally (Wagner, 2008).

* Corresponding author. Tel.: +31 20 693 0565; fax: +31 842239111.
E-mail addresses: loet@leydesdorff.net (L. Leydesdorff), wagner.911@osu.edu (C.S. Wagner), bornmann@gv.mpg.de (L. Bornmann).

http://dx.doi.org/10.1016/j.joi.2014.05.002
1751-1577/

This shift away from U.S. dominance within global publication and citation shares has attracted scholarly attention. The discussion has focused on the drop in the share percentage of the USA in scientific databases (Leydesdorff & Wagner, 2009; Shelton & Foland, 2009) and the exponential growth of Chinese contributions (Moed, 2002; Zhou & Leydesdorff, 2006). Using Web of Science (WoS, a Thomson-Reuters product) data, Leydesdorff (2012) argued that the growth of China may have been overestimated by the Royal Society (cf. Moed, Plume, Aisati, & Bervkens, 2011). An extrapolation by Shelton and Leydesdorff (2012) suggested a date beyond 2020 for the cross-over, and noted that the exponential growth of Chinese scientific publications had slowed to linear growth rates during the 2000s. The shares attributed to the European Union are less clear, partly because the borders of the EU continue to change given the accession of ten new member states in 2004, Romania and Bulgaria in 2007, and a further expansion to the EU28 most recently with the accession of Croatia in 2013.

The two multidisciplinary indexing services—Scopus and WoS—have kept pace with the growth of global science by expanding their coverage. In 2009, for example, WoS announced a regional expansion to cover more journals from Central and Eastern European countries (Testa, 2011) in response to increased coverage by Scopus, that itself was launched only in 2004. The addition of new journals was also an attempt to address the English language bias in the database (Van Leeuwen, Moed, Tijssen, Visser, & Van Raan, 2000).

A number of studies have discussed the influence of international collaborations on the global system of science (Adams, 2013; Glänzel, 2001; Luukkonen, Tijssen, Persson, & Sivertsen, 1993; Okubo, Miquel, Frigoletto, & Doré, 1992; Wagner and Leydesdorff, 2005). Multiple national addresses may partly account for the changes in shares attributed to countries (Persson, Glänzel, & Danell, 2004); but the growth of the databases and different (e.g., fractional) counting methods alone cannot account for the changes in relative positions among nations. The increased effects of networking in terms of co-authorship relations among member states of the EU has also been used as an indicator of further integration at the European level (Frenken, 2002; Frenken & Leydesdorff, 2004; Hoekman, Frenken, & Tijssen, 2010).

In this study, we address the question of whether the shifting patterns hold also for the top-1% and top-10% segments of the most-highly cited publications (that is, articles, reviews, and letters). These top-segments of the publication and citation curves represent the scientific elite, which some have argued functions as a special structure, citing one another differently from lower strata in the publication system (Cole, 1970; Mulkay, 1976). According to the results of Bornmann et al. (2010), highly cited work in all scientific fields tends to cite highly cited papers more than medium-cited work. In this study, we explore how some leading nations participate in this "elite" structure of most-highly cited publications, and whether and how this structure is influenced by and/or overlaps with international collaborations.

The *Science and Engineering Indicators* report issued by the U.S. National Science Board provides percentages of the top-1% most-highly cited publications for 2002 and 2012 (National Science Board, 2014: Appendix Table 5-57) and the numbers of publications for 13 broad fields of science and engineering in terms of six percentile rank classes (top-1%, top-5%, top-10%, top-25%, top-50%, and bottom-50% in Appendix Table 5-58; cf. Bornmann and Mutz, 2011). We take a similar approach to compare the national addresses of papers in the top-1 and top-10 percentile rank classes, but add the dynamic perspective of a decade of years (2000–2012) and include international co-authorship relations in the evaluation.

The data used in this study were harvested from an analytical version of WoS developed and maintained by the Max Planck Digital Library (MPDL, Munich). This database includes the Science Citation Index-Expanded (SCI-E), the Social Sciences Citation Index (SSCI), and the Arts and Humanities Citation Index (A&HCI) of Thomson Reuters since 1980. However, the citation impact of all papers is "field"-normalized against reference sets using the 226 WoS Categories (WC) that are attributed by Thomson Reuters to the 10,000+ journals in WoS (cf. Leydesdorff & Opthof, 2011).[1] The percentile values can be compared because they are normalized for differences among fields of science, document types, and citation windows. This organization of the data allows us to construct timelines of field-normalized impact scores (e.g., top-10%) for different nations, for groups of nations (such as the EU), and for international collaborations.

We explore the longitudinal development of the comparison between the EU28, USA, and China at the global level, and of the decomposition of the EU28 both in terms of member states and as a network of international co-authorship relations. We thus add the perspectives of using the proportions of top-1% and top-10% publications ($PP_{top\text{-}1\%}$ and $PP_{top\text{-}10\%}$; Waltman et al., 2012; cf. Tijssen, Visser, & Van Leeuwen, 2002) to the analysis in terms of percentages of world shares of publications. We did not include other nations (e.g., Japan and South Korea) in the discussion. In a follow-up study, we repeated this analysis with a focus on the BRIC(S) countries (Brazil, Russia, India, China, and South Africa; Bornmann, Wagner, & Leydesdorff, in press).

2. Methods and materials

We used integer counting to allocate publications to a country whenever this country's name is present in the publication's address lines.[2] Integer counting allows us to assume that ceteris paribus, 10% of a nation's internationally co-authored

[1] On March 3, 2014, the SCI covered 8623 journals, the SSCI 3134 journals, and the Arts & Humanities Citation Index 1727 journals. The overlap between the SCI and the SSCI is on the order of 600 journal titles.

[2] In the case of fractional counting, each country receives a fractional count based upon the number of country names in the address lines (e.g., Anderson et al., 1988; Braun et al., 1989; cf. Irvine et al., 1985). For example, if a record contains three addresses of which two are in the USA and one is in China, the record is attributed for 2/3rd to the USA and 1/3rd to China. In version 5 of WoS (since 2011), it is possible to fractionate in terms of the number of authors, since this version contains information to relate sequential authors unambiguously to address information (cf. Costas & Iribarren-Maestro, 2007; Gauffriau, Larsen, Maye, Roulin-Perriard, & von Ins, 2007).