Journal of Informetrics 5 (2011) 486–488

Contents lists available at ScienceDirect

Journal of Informetrics

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

Letter to the Editor

Some interesting insights from aggregated data published in the World Report SIR 2010

In a recently published Blog, de Moya Anegon (2011) presents aggregated bibliometric data for 50 countries worldwide having the highest paper output in the period 2004–2008 (see Fig. 1). The data for the figure are taken from the World Report SIR 2010 of the SCImago Reseach Group (2010). This report publishes quantitative data related to citation and publication "for worldwide research devoted institutions with at least 100 scientific documents published during 2008, as collected by worldwide leader scientific database Scopus." For each institution, the so called "normalized citation impact" (Karolinska Institute method, see Rehn, Kronman, & Wadskog, 2007) is reported which "reveals the ratio between the average scientific impact of an institution and the world average impact of publications of the same time period and subject area" (SCImago Reseach Group, 2010). A normalized impact of 1 means that the publications of the institution in question are (on average) no more, no less cited than an average publication in the corresponding fields of publishing.

Fig. 1 shows for each of the 50 different countries the distribution of the normalized citation impact calculated for the domestic universities. A box plot is a convenient way of visually summarizing a distribution; it consists of the lower quartile (Q1, 25% of all observations), median (50% of all observations), upper quartile (Q3, 75% of all observations), and in addition two whiskers that represent 1½ times the length of the box (Bornmann, Mutz, Neuhaus, & Daniel, 2008). Values below the lower fence or above the upper fence are outliers (Kohler & Kreuter, 2009). For example, the university with the highest impact in the USA reaches a normalized impact of approximately 2.4; the U.S. university with the lowest impact has a value of approximately 0.6. On average (median), universities in the USA have a normalized impact of 1.4 which is 40% above the

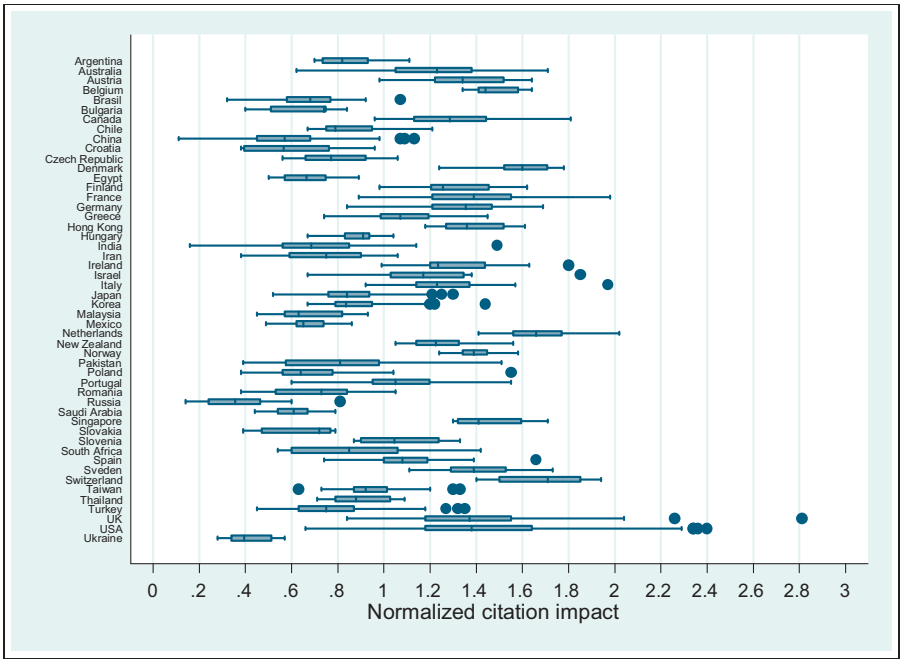

Fig. 1. Distribution of normalized impact in national systems of higher education (sorted by country).

1751-1577/$ – see front matter
doi:10.1016/j.joi.2011.03.005

Table 1
Coefficients for the correlation between normalized impact (median), number of universities within each country as well as variance, range, and maximum of normalized impact.

	Normalized impact (median)	Number of universities	Variance of normalized impact	Range of normalized impact	Maximum of normalized impact
Normalized impact (median)	1.0000				
Number of universities	−0.0930	1.0000			
Variance of normalized impact	0.2361	0.2411	1.0000		
Range of normalized impact	0.1106	0.7065[a]	0.7789[a]	1.0000	
Maximum of normalized impact	0.8780[a]	0.1861	0.5041[a]	0.4955[a]	1.0000

[a] Coefficient with Bonferroni-adjusted significance level of 0.05 or less.

world average of 1. The box plots shown in the figure point out interesting differences between the countries (see here de Moya Anegon, 2011). Nearly half of the countries have more than 75% of its universities below and the other half above the world average. Above the average are countries like USA, UK, and Germany, and below are for example China, Japan, India, and Russia. Furthermore, there are large differences visible between the countries concerning the range or variance, respectively, of the institutional normalized impact. Whereas the impact values are very heterogeneous for the USA and UK, they are homogeneous for Belgium and the Netherlands.

We calculated correlations (Spearman's rank correlations) between some aggregated key figures based on the institutional normalized impact values which are shown in Fig. 1. The correlation coefficients may provide information on questions like: What drives the overall research performance of a country measured by the median normalized impact? In Table 1, coefficients for the correlation between the number of universities within each country, median normalized impact, as well as variance, range, and maximum of normalized impact are shown. Coefficients that are statistically significantly different from 0 are marked. The strength of the relationships between the different variables in the table is interpreted by using the guidelines of Cohen (1988) and Kraemer et al. (2003).

As the coefficients show for the median normalized impact, the overall performance of a country does scarcely correlate with the numbers of the universities and the range/variance of the normalized impact. The coefficients ($r = -0.093$, $r = 0.2361$, $r = 0.1106$) are not statistically significant and in the range of small or smaller than typical in the applied behavioural sciences. That means not only larger industrial nations with a lot of universities (e.g., the USA) have a chance of a high performance but also smaller countries (e.g., Switzerland). Furthermore, the country's overall impact is not dependent on the heterogeneity or homogeneity, respectively, of the universities' performances within the country. In many countries, for example Germany, there is a trend to a more heterogeneous higher education system with some excellent universities at the top to strengthen the position of the country in total against other countries. However, the relatively low coefficients found in this study challenge this effort of enhancing heterogeneity.

It is interesting to see in Table 1 that the median normalized impact is correlated much larger than typical ($r = .878$) with the maximum of the normalized impact. The best university within a country seems to determine the country's average impact (and vice versa). This result is not an artefact of the measure of central tendency for impact; the here used median is scarcely affected by extreme values in contrast to the arithmetic average. Either the universities within a country profit from the best university or the best university within a country cannot be much better than the national standard. Table 1 further shows that the maximum of impact is correlated only smaller than typical with the number of universities ($r = .186$): To have many universities within one country does not automatically result in a high-performing institution, although it seems actually reasonable that more competition among the universities within one country would lead to at least one high-performing university on the world level.

In this Letter to the Editor, some interesting insights into the research impact of universities in different countries are given. One should be aware that correlations cannot give explanations for a phenomenon. "Regardless of the strengths of correlation, there is no necessarily a causal relationship between the variables" (Kohler & Kreuter, 2009, p. 183). They can only give stimuli to think about possible reasons for a relationship.

References

Bornmann, L., Mutz, R., Neuhaus, C., & Daniel, H.-D. (2008). Use of citation counts for research evaluation: Standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics*, *8*, 93–102. doi:10.3354/esep00084

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ, USA: Lawrence Erlbaum Associates.

de Moya Anegon, F. (2011). The research impact of national higher education systems. Retrieved March 2, from http://scimagolab.com/blog/2011/the-research-impact-of-national-higher-education-systems/.

Kohler, U., & Kreuter, F. (2009). *Data analysis using stata* (2nd ed.). College Station, TX, USA: Stata Press, Stata Corporation.

Kraemer, H. C., Morgan, G. A., Leech, N. L., Gliner, J. A., Vaske, J. J., & Harmon, R. J. (2003). Measures of clinical significance. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(12), 1524–1529. doi:10.1097/01.chi.0000091507.46853.d1

Rehn, C., Kronman, U., & Wadskog, D. (2007). *Bibliometric indicators—Definitions and usage at Karolinska institutet*. Stickholm, Sweden: Karolinska Institutet University Library.

SCImago Reseach Group. (2010). *SIR World Report 2010: Global Ranking*. Granada, Spain: University of Granada.

Lutz Bornmann*
Max Planck Society, Hofgartenstr. 8, D-80539 Munich, Germany

Felix de Moya Anegon
CSIC/CCHS/IPP, Albasanz 26, Madrid, Spain

* Corresponding author.
E-mail address: bornmann@gv.mpg.de (L. Bornmann)

29 March 2011