



Contents lists available at ScienceDirect

Journal of Informetrics

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

Letter to the Editor

Is it necessary to consider suburbs (or small cities in the close proximity) and name variants in a citation impact analysis for bigger cities? An investigation using Munich as an example

In a recently published *Nature* report van Noorden (2010) presented some studies and Web applications for the comparison of bigger cities worldwide with regard to publication output and citation impact (see here also Frenken, Hardeman, & Hoekman, 2009). Results like the following are presented from the studies: “Beijing, which churned out 0.76% of the global output in 1996, produced 2.74% in 2008 (319,000 research papers). Other fast-growing areas are Tehran, Istanbul, Seoul, Singapore City and São Paulo” (van Noorden, 2010, p. 907). However, the results of the different studies and Web applications concerning output and impact are not directly comparable: How one scientist defines a city area is not consistent with how another would do this (and it depends especially on familiarity with the city and institutes involved). For example, Matthiessen, Schwarz, and Find (2010, p. 1882) have combined neighbouring cities if the cities are reachable with a “less than 45 minutes of surface transport”. In the Web applications presented as online collection by van Noorden (2010) the citation impact and publication output of each city is analysed without considering neighbouring cities. Thus, the question arises of city performance differences in dependency on the grouping approaches of the scientists: does it make a difference to consider (1) only the main city variant used by authors of publications (without considering neighbouring cities) or (2) all city name variants in addition to cities in proximity? We investigated this question using Munich (in German: München) as an example. The results of our analyses are presented in Table 1. The data come from a static version of the Scopus database extracted in 2009. We restricted not only to articles, reviews and conference papers but also to a “core” Scopus journal set which excludes stand-alone conference proceedings or very patchily covered journals.

The table shows article counts and relative citation rates of articles published with certain city name variants used by authors. Besides München, Munchen (without umlaut) is also indexed as Munich; Garching, Oberschleissheim and Martinsried are suburbs of Munich (or small cities closed to Munich). Publications for a given city variant (for example, “Garching”) involves no double-counting – all publications that contain the term “Garching” in the AFFILCITY field in the underlying metadata are counted here, and not in any other variant’s count. So, the data shown pertains solely to that variant in the dataset. The relative citation rates shown for the different name variants in the table are journal-oriented field-normalized citation ratios: each city’s observed citation count is divided by the expected citation count for the same document types published within the same publication year in the same field(s) (defined by a journal-level field classification). Where a city has published documents in journals classified in more than one field, an average of the expected citation counts in these fields is used to derive the expected citation count. A relative citation rate of 1 means that the city in question is no more and no less cited than expected for an “average” city publishing in the corresponding field(s).

The numbers for the name variant München in Table 1 are exactly the same numbers like those which can be extracted from the Web application “Output versus Impact” on <http://www.nature.com/news/specials/cities/best-cities.html>. This variant incorporates clearly more articles than all other variants: For example, whereas München has 34,260 articles in the year 2000, Oberschleissheim has only 258. Thus München is the dominant variant for Munich and its close proximity which should be compared with all other different variants (see the notes of the table). The relative citation rates for all different variants are shown in the line total. As the small differences between München and total in Table 1 reveals the city name variants do not add much citation impact to München. The differences are between 0.07 (year: 2006) and 0.01 (year: 2001).

These results indicate that for Munich the citation impact values do not depend on the consideration of suburbs or small cities, respectively, in close proximity and different name variants besides the variant München. Since our analyses refer to only one city, our results are not generalizable. They furnish first evidence which should be proofed with other cities. This Letter to the Editor is intended to motivate other researchers to undertake these investigations and to develop standards in this respect. Citation impact analyses for city areas are often done in a more or less subjective way.

Table 1
Number of articles and relative citation rates (impact) for articles published with different city name variants of Munich and its suburbs.

City name variant	Publication year																	
	2000		2001		2002		2003		2004		2005		2006		2007		2008	
	Articles	Impact	Articles	Impact	Articles	Impact	Articles	Impact	Articles	Impact	Articles	Impact	Articles	Impact	Articles	Impact	Articles	Impact
München	34,260	1.39	34,130	1.42	33,410	1.45	34,265	1.47	36,370	1.46	39,112	1.46	42,066	1.48	46,284	1.50	48,653	1.53
Garching bei München	4959	1.82	5139	1.81	5185	1.86	5376	1.87	5872	1.97	6397	2.06	6991	2.20	7566	2.21	7965	2.16
Munich	4666	1.03	4531	1.09	4310	1.12	4420	1.11	4654	1.14	5010	1.14	5364	1.12	5907	1.09	6035	1.10
Garching	280	1.74	366	1.58	421	1.54	481	1.48	494	1.43	570	1.63	539	1.66	606	1.84	581	1.79
Garching bei München	306	1.35	340	1.34	349	1.42	374	1.48	409	1.46	449	1.49	468	1.43	488	1.28	506	1.44
Oberschleissheim	258	1.55	248	1.61	245	1.53	228	1.29	219	1.37	242	1.32	278	1.36	309	1.45	442	1.39
Martinsried	29	2.81	29	2.03	30	1.91	27	1.34	22	1.48	23	1.24	23	1.08	26	1.30	27	1.69
Total	44,758	1.41	44,783	1.43	43,951	1.47	45,172	1.49	48,041	1.49	51,804	1.51	55,730	1.55	61,186	1.56	64,209	1.58
Difference between total and München		0.02		0.01		0.02		0.02		0.03		0.05		0.07		0.06		0.05

Notes: Owing primarily to continuous improvements in the affiliation matching algorithms in Scopus, expected to reach a steady-state level of high precision and recall in 2011, the data presented here for some of the (relatively) minor city names may appear lower than expected. For example, the data here show just 27 articles attributed to authors in Martinsried (home of the Max Planck Institute of Biochemistry, among other institutes) in the 5 years ending 2008, which can be shown to be almost an order of magnitude lower than the count from a comparable search at www.Scopus.com. The shortfall is likely to have been attributed instead to other variants of these affiliations, and thus counted under Munich in the current data. In addition to the improvements in affiliation matching as discussed above, a portion of the difference may be attributed to articles being added to the database, or having missing affiliation data added, since the current data were extracted. However, to allow for direct comparison with the data as published in the *Nature* article, and for the computation of field-weighted citation impact (which is not directly possible in Scopus itself) we have elected to present these numbers as they are.

References

- Frenken, K., Hardeman, S., & Hoekman, J. (2009). Spatial scientometrics: Towards a cumulative research program. *Journal of Informetrics*, 3(3), 222–232. doi:10.1016/j.joi.2009.03.005
- Matthiessen, C. W., Schwarz, A. W., & Find, S. (2010). World cities of scientific knowledge: Systems, networks and potential dynamics. An analysis based on bibliometric indicators. *Urban Studies*, 47(9), 1879–1897. doi:10.1177/0042098010372683
- van Noorden, R. (2010). Cities: Building the best cities for science. Which urban regions produce the best research—And can their success be replicated? *Nature*, 467, 906–908. doi:10.1038/467906a

Lutz Bornmann*

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

Andrew Plume

Elsevier, The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

*Corresponding author.

E-mail address: bornmann@gv.mpg.de

(L. Bornmann)

11 May 2011