

The Effect of Having a Research Chair on Scientists' Productivity

Seyed Reza Mirnezami¹ and Catherine Beaudry²

¹*seyed-reza.mirnezami@polymtl.ca* ²*catherine.beaudry@polymtl.ca*
Polytechnique Montreal, P.O. Box. 6079, Montreal, QC, H3C 3A7 (Canada)

Abstract

Having combined data on Quebec scientists' funding and journal publication, this paper tests the effect of having a research chair on the scientists' performance. The novelty of this paper is to use matching technique to understand whether having a research chair is a real cause for better scientific performance. This method compares two different sets of regressions, which are conducted on different data sets: the one with all records and another with records of matched scientists only. Two chair and non-chair scientists are called matched with each other when they have closest propensity score in terms of age, number of articles, and amount of funding. The result shows that research chair is a significant determinant in complete data set but it is insignificant when only matched scientists are kept in data set. In other words, in the case of two scientists with similarity in terms of three mentioned factors, having a chair cannot significantly affect the scientific performance.

Conference Topic

Science policy and research assessment

Introduction

The scientists' academic performance has been extensively discussed and many of its determinants are currently known as potential motives for publishing papers in peer reviewed journals. Among others, age, gender, private and public funding, institutional setting, field and context are the most important determinants. The funding definitely plays the major role in knowledge production and shaping scientific productivity. Its positive effect has been extensively investigated in literature (Crespi & Geuna, 2008; Pavitt, 2000, 2001; Salter & Martin, 2001).

However, having a great academic performance does not depend solely on funding. The networking capability of scientist can also explain the number of journal papers. Most of the studies on the effects of network rely on co-authorship as a proxy of scientific collaboration (Katz & Martin, 1997; Melin & Persson, 1996). In addition to direct collaboration, there are also some other networking measures, which are known in the literature as determinant of publication. For instance, it is possible to show how a researcher links two other researchers by making separate collaborations with them. Newman (2001a, 2001b) finds that in physics, biomedical research, and computer science, most of the authors are connected with each other via one or two of their collaborators, a concept generally referred to as betweenness centrality. Beaudry and Allaoui (2012) also show a positive effect of betweenness centrality on the scientific productivity of Quebec's scientists.

In addition to the above measures of networking effect, the networking capacity of scientists partially depends on prestige of their academic affiliation. Turner and Mairesse (2005) show it for the outstanding performance of 'Grandes Ecoles' in France. Beside the name and brand of academic institutions, centers with specific research orientations such as 'centers for excellence' are also effective. According to Niosi (2002), the government of Canada launched 7 centers for biotechnology sectors in 1988, which financially supports the collaboration of university research, the specialized biotechnology firms, and the governmental laboratories. In addition to the funding support, however, this program comes up with improving intellectual property regulations, and developing human resources.

There are some other desirable factors similar to ‘centers for excellence’, which increase an individual’s research motivation and influence the willingness or ability of scientists for conducting original research. In this paper, we focus on the effect having a ‘research chair’ as a possible determinant of scientific publication. On the one hand, it helps the holder of this position to absorb more money or to construct more effective network, which results in propelling future knowledge production. On the other hand, it may be the effect of past super performance of scientist, implying the intrinsic ability of scientists in conducting research or referring to the chair-holder extensive networking capacity.

By analysing data in an econometric model, it is possible to test the significant effect of ‘being a chair holder’ on the scientific productivity. The rest of paper is followings: Section 1 reviews theoretical framework and literature review. Section 2 explains how data is gathered and what the variables represent. In addition, it raises the related hypotheses and explains which econometric models can test these hypotheses. Section 3 presents the results of econometric model and the result of testing hypotheses. A conclusion will summarize the results of the paper.

Section 1 - Theoretical framework

The literature relevant to this article brushes on the importance of having a prestigious academic position or affiliation. Focusing on the role of university prestige in academic performance, Long, Allison, and McGinnis (1979) found a positive and significant correlation between the prestige of the scientist alma matter and prestige of subsequent employment affiliation. The authors also indicated that graduating from a prestigious university has a positive effect on citations (but not on publication counts). The paper also provides a justification for the effect of prestige arguing that the best students are admitted to the most prestigious universities and subsequently the graduates of the prestigious universities are generally recruited by other similar institutions. Furthermore, such scientists who studied in and have been recruited by prestigious universities are better able to interact with new gifted students (Long et al., 1979). This paper tries to argue that academic prestige can push forward research and its quality. More recently, Zhou, Lü, and Li (2012) show that papers cited by prestigious scientists, regardless of the number of citations, are of a higher quality than papers which are cited by ‘ordinary’ scientists.

The prestige can be seen from the reverse direction of causality. West, Smith, Feng, and Lawthom (1998) investigate the relationship between departmental climate, such as degree of formalization, support for career development and support for innovation on the one hand, and official rated effectiveness of universities on the other hand. They conclude that the causality direction is from former to latter, showing that prestige of universities is an effect and not a cause for appropriate departmental climate and necessary institutional setting for conducting research.

Nevertheless, measuring academic prestige itself is another story. Frey and Rost (2010) compare three types of university ranking based on the number of articles, number of citations, and membership of editorial board or of academic associations. The paper indicates that these rankings are not compatible with each other and suggests the use of multiple measurements. Van Raan (2005) criticize the applicability of university rankings such as the Shanghai ranking for evaluating academic excellence by noting that the ‘affiliation’, as an important factor reflecting research atmosphere, is not well addressed in those ranking. In addition to the university ranking, it is important to assess individual research productivity to have a better sense of prestige. Henrekson and Waldenström (2007) introduce three types of indicators, measuring research performance: (1) measures based on weighted journal publications, (2) measures based on citations to most cited works, and (3) measures based on the number of publications.

To measure prestige with more robust measure, it is possible to consider the honor as the measure of prestige, which is awarded based on a deliberate assessment in specialized and independent committees. Different types of research chair are example of awards. In Canada, there are three types of research chair: (1) the research chairs which are awarded by industry and called industrial chair; (2) the research chairs which are awarded by Canadian funding agencies such as NSERC, SSHRC, and CIHR; and (3) the ‘Canada research chairs’, whose holders are assumed to already achieve research excellence in one main fields of research: engineering and the natural sciences, health sciences, humanities, and social sciences. The purpose of this program is to “improve our depth of knowledge and quality of life, strengthen Canada's international competitiveness, and help train the next generation of highly skilled people through student supervision, teaching, and the coordination of other researchers’ work”.¹ Considering this specific measure of prestige, it is possible to find out the effect of being a ‘chair-holder’ on scientific productivity. Therefore, our first hypothesis reads as:

Hypothesis 1

Being chair-holder increases the scientist’s number of publications.

The hypothesis 1 just tests the performance of chair-holders compared to other scientists and it does not seek for the cause and effect. Considering the fact that the chair-holders are the well-funded scientists too, this hypothesis cannot detach the funding effect of chair from its other effects (mainly from prestige and networking effect). In other words, there are evidences in literature about the benefits and goals of research chair program other than funding, but hypothesis 1 is not able to test them.

Some articles try to highlight the functions and characteristics of research chair. Cantu, Bustani, Molina, and Moreira (2009) show the research chair program would be a good strategy for implementing knowledge-based development. In study on German universities, Schimank (2005) argues that chair-holders are small businessmen with high job security and no bankruptcy in addition to the good level of freedom of teaching and research, indicating that research chair has characteristics of job security and sovereignty.

According to some official documents, affecting scientific productivity is not the direct goal of research chair. In the tenth-year evaluation report for Canada research chair (CRC),² the authors conclude that CRC program is an effective way for Canadian universities to “attract and retain leading researchers” from other countries. The report does not say that having a research chair is determinant and cause of chair’s scientific production: “the extent to which this success can be related directly to the CRC is difficult to quantify”. It is also possible to bring some evidence that having a research chair is not a cause for other factors such as salary. Courty and Sim (2012) show that although having Canada Research Chair (CRC) initially increases the professors’ salary, such increase erodes quickly over the time. This means that getting a research chair does not necessarily result in long term salary jump.

Regarding the mentioned points, it is possible to look at the research chair as the effect of scientists’ characteristics (including age, number of articles, and number of citations), while it aims to expand academic network and absorb highly skilled talents. To control for the effect of scientist’s past performance on having a research chair and to detach the funding advantage of chair, we propose our second hypothesis:

Hypothesis 2

Keeping the main scientists’ characteristics (age, number of articles, and amount of grant) constant, having a research chair does not have significant positive effect on scientists’ productivity.

¹ http://www.chairs-chaires.gc.ca/about_us-a_notre_sujet/index-eng.aspx

² http://www.chairs-chaires.gc.ca/about_us-a_notre_sujet/publications/ten_year_evaluation_e.pdf

This hypothesis can be tested by matching technique, which will be explained in the methodology section. The important note here is that ‘being a research chair’ cannot be the only determinant in right-hand-side of regression equations. We should look for some control variables, which are mentioned in literature as determinants of scientific production. Among others, age, gender, funding, field, and university characteristics are the most important determinants of scientific production which should be controlled when the effect of research chair on scientific productivity are being tested.

In terms of age, there are two groups of evidences in literature about its effect on scientific productivity. First, some articles assess the life cycle trend in economic activity, referring to the non-linearity of human productivity during life (Becker, 1962). The second group of articles generally find that scientists’ academic performance (number of articles and number of citations) decreases as they age (Bonaccorsi & Daraio, 2003; Diamond, 1986; Levin & Stephan, 1991). Some articles like Gonzalez-Brambila and Veloso (2007) also indicate that age does not have any effect on the number of articles but it positively affect the number of citations. Gender effect is known as a significant determinant of scientific productivity in literature. Long (1990) explains that women’s opportunities for collaboration are significantly less than those of men’s because women have young children. However, in another study, Long (1992) shows that women are less productive in the first decade of their career but are more productive afterwards. Research funding is another important determinant of scientific productivity. Pavitt (2001) also refers to the importance of public support for scientific infrastructure development and highlights its role in the effectiveness of public grants. In another study, Pavitt (2000) argues that funding for infrastructure of expertise, equipment and networks is necessary for development and implementation of research. A body of literature investigates the effect of university characteristics on the scientific productivity. There are also some papers about the effect of faculty size. Buchmueller, Dominitz, and Lee Hansen (1999) indicate that graduate school faculty size is a significant determinant of the research proficiency of graduates. Jordan, Meador, and Walters (1988, 1989) indicate that research productivity is positively associated with department size but that effect becomes weaker as the size increases. In an opposite direction, Kyvik (1995) rejects both hypotheses that large departments are more productive and that faculty members of large departments better assess the research environment.

There also some evidences about differences between fields and context. Blackburn, Behymer, and Hall (1978) show that the fields of humanities and sciences have different pattern of scientific production. To justify the differences between disciplines, Baird (1986) shows that for instance large research laboratory in chemistry, scholarly apprenticeship approach in history, and research over practice in psychology are important factors in scientists’ productivity, which are field-dependent factors. In another comprehensive study, Baird (1991) refers to the productivity and citation pattern differences among disciplines and argues that size, internal university support and federal support can explain such differences. All of the mentioned evidence in literature shows that scientific productivity may have different determinants including academic prestige and other control variables such as funding, gender, age, and university-specific characteristics.

Section 2 - Data and methodology

Data and variables

In order to validate these two hypotheses, we built a data set based on the integration of data on funding and journal publications for Quebec scientists. For publications, Elsevier’s Scopus provides information on scientific articles (date of publication, journal name, authors and their affiliations). In terms of funding, there is a database for researchers in Quebec universities

(Système d'information sur la recherche universitaire or SIRU) gathered and combined by the Ministry of Education, Leisure and Sports. The SIRU database lists the grants and contracts information, including yearly amount, source, and type during the period of 2000-2010 for all Quebec university scientists. The appendix 1 reviews the names and description of variables in data set.

Methodology and econometrics model

To measure the effect of 'being a research chair' on the scientist's performance, a regression equation is fitted to the available data using a panel regression. In such regression, the left-hand-side (LHS) variable of regression is the number of articles [$\ln(nbArticle)$] as a measure of scientific productivity. In terms of right-hand-side (RHS) variable, the main independent variables are the dummy variables of research chair [$dChair1$, $dChair2$, $dChair3$, $dChair4$, $dChair5$]. However, the dependent variable of regression in LHS should be also controlled for the other determinants of articles count. Among others, age [Age], gender [$dFemale$], and funding are the important ones. We also control for the fixed effect of university, year, and research field in order to account for any impact that our explanatory variables do not cover.

It is important to note that two variables of [$\ln(PublicfundingO)$] and [$\ln(nbArticle)$] are determined by each other and co-evolved during time, which is the source of endogeneity. Thus it means that simple ordinary least square or panel models are biased. The main reason for this potential endogeneity is that scientists are assessed for public funding based on their CV and past performance while at the same time, publication and research quality significantly depends on the funding capability of researchers. Using instrumental variables (IV) instead of endogenous variable is a common suggested method in literature to address endogeneity problem. If there is more than one instrument for an endogenous variable, it is necessary to perform a two-stage regression, in which the first stage estimates the endogenous variable (named here as instrumented variable) based on a list of instrumental variables. In the first stage of our model, the amount of public funding [$\ln(PublicfundingO)$] is estimated by the rank of scientist in the field in terms of three-year average of funding (for the purpose of operational costs and direct expenditure of research) [$PubORank$], the rank of scientist in the field in terms of three-year average of articles count [$PublRank$], and natural logarithm of three-year average of aggregate public sector funding in the field [$\ln(totFund)$]. These three variables play the role of instruments for public funding. It should also be noted that public funding is not determined by the instruments in the same year. Hence the one-year lags of instruments are being used in the first-stage regression. The second stage is similar to the previous model in which there is no endogeneity.

1st stage: $\ln(PublicfundingO)_{it} = f(PubORank_{it-1}, PublRank_{it-1}, \ln(totFund)_{it-1})$

2nd stage: $\ln(nbCitation)_{it} = f(\ln(PublicfundingO)_{it}, \ln(PrivatefundingO)_{it}, \ln(NFPfundingO)_{it}, (dChair1|dChair2|dChair3|dChair4|dChair5)_{it}, dFemale, Age, Age^2, \text{research field dummies, year dummies, university dummies})$

The main purpose of this research is to show how much having a research chair as an external support is important and significant in promoting scientific publication. To test the first hypothesis, it is sufficient to run the two-stage panel regression on the whole data set whether 'having a research chair' is a significant RHS variable, either as a real cause or a channel for other variables/causes. According to the chair characteristics, the networking and prestige effect of 'having a research chair' may be mixed with the effect of funding. To address this issue, we use matching technique and compare two chair and non-chair scientists who have close funding to each other (and have some other similar characteristics). Like what Bérubé and Mohnen (2009) did, it is possible to find pairs of chair and non-chair by using the `psmatch2` command in Stata and delete the unmatched records. The selection is made by

generating propensity score and choosing the pairs of scientists with closest scores to each other. The new data set consists of twin scientists who are similar to each other in terms of funding, gender, and division of studies.³

By controlling the mentioned criteria and keeping matched scientists only, ‘having a research chair’ becomes a better and more informative signal for the prestige and networking of scientists. In this case, the effect of ‘being chair’ on scientific productivity does not include funding effect or it is not related to the division or gender of scientist. To test the second hypothesis, only matched pairs of scientists are being used in regression analysis to identify whether having a research chair is a significant cause for scientific productivity.

One of the important stages in matching technique is to check the quality of matching. It means there should be no difference between the averages of mentioned criteria (gender, funding, and division of studies) when the comparison is made between chair and non-chair scientists among the matched pairs. However, there can be a difference when the comparison is made in original database and before any entry deletion. Table 1 summarizes such comparisons to show that the matching is done with an acceptable quality for *dChair3*, *dChair4*, and *dChair5*.

Table 1. Make a comparison between mean to show the quality of matching.

	Comparison over whole database				Comparison over matched scientists “After Matching”			
	Gender	Funding	Research field ⁴	number of scientist	Gender	Funding	Research field	number of scientist
<i>dChair3</i> =0	0.2959	86217	0.4284	7359	0.1023	403051	0.2286	293
<i>dChair3</i> =1	0.2013	464106	0.3447	293	0.2013	464106	0.3447	293
Is difference significant at 5% level?	Yes	Yes	Yes		Yes	No	No	
<i>dChair4</i> =0	0.2954	95871	0.4318	7508	0.1111	369080	0.0416	144
<i>dChair4</i> =1	0.1319	351785	0.0833	144	0.1319	351785	0.0833	144
Is difference significant at 5% level?	Yes	Yes	Yes		No	No	No	
<i>dChair5</i> =0	0.2987	82183	0.4344	7234	0.1483	367494	0.1698	418
<i>dChair5</i> =1	0.1818	420920	0.2655	418	0.1818	420920	0.2655	418
Is difference significant at 5% level?	Yes	Yes	Yes		No	No	No	

Section 3 - Result and discussion

Based on the models presented in methodology section, we need to first run the regressions on the whole dataset (Table 2) which show that all types of chair have positive and significant effect on scientific productivity. However after keeping only matched scientists in dataset, who are similar to each other in terms of gender, funding, and research field, the regression equations indicate significant and positive result only for Canada research chair (Table 3) Industrial chairs and chairs appointed by Canada research council (NSER, SSHRC, and CIHR) do not have an independent positive effect on scientific productivity. Considering the hypotheses in previous section, it possible to validate the first hypothesis and partially validate the second hypothesis. One may question whether research chairs in general are independent cause for research productivity or they are proxy for other known factors in literature. Considering literature and mentioned mission of research chairs in their mandate,

³ We have three divisions: ‘engineering and the natural sciences’, ‘health sciences’, and ‘humanities, and social sciences’

⁴ Test whether dummy variable of Social Science and Humanities is equal to 1.

Table 2. Regression results over all samples for *dChair3* and *dChair4* (the second stage of 2SLS).¹

$\ln(nbArticle)_{it}$	IV1	IV2	IV3	IV4	IV5	IV6	IV7	IV8	IV9	IV10	IV11
$\ln(PublicfundingO)_{it}$	0.0433 ***	0.0417 ***	0.0417 ***	0.0416 ***	0.0417 ***	0.0416 ***	0.0415 ***	0.0415 ***	0.0417 ***	0.0416 ***	0.0416 ***
	0.0011	0.0011	0.0011	0.0011	0.0011	0.0011	0.0011	0.0011	0.0011	0.0011	0.0011
$\ln(PrivatefundingO)_{it}$	0.0112 ***	0.0109 ***	0.0105 ***	0.0109 ***	0.0105 ***	0.0113 ***	0.0108 ***	0.0111 ***	0.0110 ***	0.0109 ***	0.0110 ***
	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007
$\ln(NFPfundingO)_{it}$	0.0076 ***	0.0074 ***	0.0074 ***	0.0075 ***	0.0075 ***	0.0074 ***	0.0092 ***	0.0092 ***	0.0074 ***	0.0075 ***	0.0074 ***
	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007	0.0006	0.0006	0.0006
Age_{it}	0.0021	0.0038	0.0038	0.0038	0.0038	0.0037	0.0036	0.0036	0.0038	0.0038	0.0038
	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025
sq_Age_{it}	-0.0001 ***	-0.0001 ***	-0.0001 ***	-0.0001 ***	-0.0001 ***	-0.0001 ***	-0.0001 ***	-0.0001 ***	-0.0001 ***	-0.0001 ***	-0.0001 ***
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$dFemale_i$	-0.0911 ***	-0.0847 ***	-0.0848 ***	-0.0847 ***	-0.0848 ***	-0.0815 ***	-0.0700 ***	-0.0686 ***	-0.0832 ***	-0.0841 ***	-0.0827 ***
	0.0109	0.0108	0.0108	0.0108	0.0108	0.0110	0.0112	0.0113	0.0109	0.0109	0.0109
$dFemale_i * \ln(PrivatefundingO)_{it}$						-0.0023		-0.0013			
						0.0016		0.0016			
$dFemale_i * \ln(NFPfundingO)_{it}$							-0.0065 ***	-0.0064 ***			
							0.0013	0.0013			
$dChair3_{it}$		0.3331 ***	0.3105 ***	0.3444 ***	0.3233 ***	0.3332 ***	0.3323 ***	0.3324 ***	0.3330 ***	0.3413 ***	0.3404 ***
		0.0249	0.0268	0.0271	0.0284	0.0249	0.0249	0.0249	0.0251	0.0252	0.0254
$dChair4_{it}$		0.1025 ***	0.1006 **	0.0891 **	0.0894 **	0.1020 ***	0.0998 ***	0.0996 ***	0.1195 ***	0.0942 ***	0.1114 ***
		0.0352	0.0432	0.0387	0.0451	0.0352	0.0352	0.0352	0.0360	0.0356	0.0362
$dChair3_{it} * \ln(PrivatefundingO)_{it}$			0.0060 **		0.0064 **						
			0.0026		0.0027						
$dChair4_{it} * \ln(PrivatefundingO)_{it}$			0.0000		-0.0003						
			0.0033		0.0034						
$dChair3_{it} * \ln(NFPfundingO)_{it}$				-0.0026	-0.0033						
				0.0024	0.0024						
$dChair4_{it} * \ln(NFPfundingO)_{it}$				0.0026	0.0026						
				0.0031	0.0031						
$dFemale_i * \ln(PrivatefundingO)_{it} * dChair3_{it}$									0.0005		0.0024
									0.0063		0.0064
$dFemale_i * \ln(PrivatefundingO)_{it} * dChair4_{it}$									-0.0177 **		-0.0212 **
									0.0077		0.0079
$dFemale_i * \ln(NFPfundingO)_{it} * dChair3_{it}$										-0.0102 **	-0.0104 **
										0.0050	0.0050
$dFemale_i * \ln(NFPfundingO)_{it} * dChair4_{it}$										0.0125	0.0175 **

¹ *, **, and *** show the significance level at 0.05, 0.02, and 0.01 respectively - Year dummies, field dummies, and university dummies are significant. The minimum year activity, average year activity, and maximum year activity are 1, 10.6, and 12 respectively.

$\ln(nbArticle)_{it}$	IV1	IV2	IV3	IV4	IV5	IV6	IV7	IV8	IV9	IV10	IV11
<i>Constant term</i>	0.4681 ***	0.4210 ***	0.4218 ***	0.4200 ***	0.4204 ***	0.4222 ***	0.4218 ***	0.4223 ***	0.4202 ***	0.4210 ***	0.4205 ***
<i>Number of observations</i>	0.0683	0.0680	0.0680	0.0680	0.0680	0.0680	0.0680	0.0680	0.0680	0.0680	0.0680
<i>Number of scientists</i>	80772	80772	80772	80772	80772	80772	80772	80772	80772	80772	80772
χ^2	7652	7652	7652	7652	7652	7652	7652	7652	7652	7652	7652
<i>sigma</i>	13859.3	14234.6	14251.6	14244.3	14258.4	14236.5	14277.1	14277.9	14239.7	14241.7	14246.4
<i>rho</i>	0.5689	0.5664	0.5661	0.5662	0.5660	0.5664	0.5662	0.5662	0.5664	0.5664	0.5664
R^2 within groups	0.4235	0.4183	0.4178	0.4180	0.4176	0.4184	0.4181	0.4182	0.4183	0.4183	0.4184
R^2 overall	0.0617	0.0630	0.0629	0.0631	0.0630	0.0631	0.0633	0.0634	0.0631	0.0631	0.0632
R^2 between groups	0.3367	0.3456	0.3460	0.3455	0.3458	0.3457	0.3464	0.3464	0.3457	0.3456	0.3457
	0.5044	0.5148	0.5154	0.5145	0.5151	0.5148	0.5156	0.5156	0.5148	0.5147	0.5148

Table 3. Regression results over only matched pairs of scientists for *dChair3* and *dChair4* (the second stage of 2SLS).²

$\ln(nbArticle)_{it}$	IV23	IV24	IV25	IV26	IV27	IV28	IV29	IV30	IV31	IV32	IV33
$\ln(PublicfundingO)_{it}$	0.0702 ***	0.0680 ***	0.0692 ***	0.0680 ***	0.0691 ***	0.0680 ***	0.0679 ***	0.0679 ***	0.0682 ***	0.0678 ***	0.0679 ***
$\ln(PrivatefundingO)_{it}$	0.0059	0.0059	0.0060	0.0059	0.0060	0.0059	0.0059	0.0059	0.0059	0.0059	0.0059
$\ln(NFPfundingO)_{it}$	0.0053 ***	0.0059 ***	0.0076 ***	0.0059 ***	0.0072 ***	0.0062 ***	0.0059 ***	0.0062 ***	0.0066 ***	0.0060 ***	0.0067 ***
Age_{it}	0.0019	0.0019	0.0025	0.0019	0.0026	0.0021	0.0019	0.0021	0.0020	0.0019	0.0020
sq_Age_{it}	0.0038 **	0.0042 **	0.0041 **	0.0077 **	0.0074 **	0.0041 **	0.0045 **	0.0044 **	0.0041 **	0.0045 **	0.0043 **
$dFemale_i$	0.0018	0.0018	0.0018	0.0025	0.0026	0.0018	0.0019	0.0019	0.0018	0.0019	0.0019
$dFemale_i * \ln(PrivatefundingO)_{it}$	0.0217 **	0.0244 **	0.0260 **	0.0249 **	0.0265 **	0.0244 **	0.0244 **	0.0244 **	0.0246 **	0.0242 **	0.0243 **
$dFemale_i * \ln(NFPfundingO)_{it}$	0.0104	0.0104	0.0104	0.0104	0.0104	0.0104	0.0104	0.0104	0.0104	0.0104	0.0104
$dChair3_{it}$	-0.0003 ***	-0.0003 ***	-0.0003 ***	-0.0003 ***	-0.0003 ***	-0.0003 ***	-0.0003 ***	-0.0003 ***	-0.0003 ***	-0.0003 ***	-0.0003 ***
	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
	-0.1217 **	-0.1230 **	-0.1215 **	-0.1224 **	-0.1210 **	-0.1160 **	-0.1138 **	-0.1081 **	-0.0889 **	-0.1173 **	-0.0848 **
	0.0545	0.0533	0.0533	0.0532	0.0533	0.0562	0.0572	0.0595	0.0555	0.0549	0.0567
						-0.0020		-0.0018			
						0.0051		0.0051			
							-0.0022	-0.0020			
							0.0049	0.0049			
		0.1696 ***	0.1625 ***	0.2062 ***	0.1954 ***	0.1697 ***	0.1696 ***	0.1698 ***	0.1689 ***	0.1766 ***	0.1756 ***
		0.0451	0.0483	0.0483	0.0506	0.0451	0.0451	0.0452	0.0453	0.0454	0.0456

² *, **, and *** show the significance level at 0.05, 0.02, and 0.01 respectively - Year dummies, field dummies, and university dummies are significant. The minimum year activity, average year activity, and maximum year activity are 1, 10.9, and 12 respectively.

$\ln(nbArticle)_{it}$	<i>IV23</i>	<i>IV24</i>	<i>IV25</i>	<i>IV26</i>	<i>IV27</i>	<i>IV28</i>	<i>IV29</i>	<i>IV30</i>	<i>IV31</i>	<i>IV32</i>	<i>IV33</i>
<i>dChair4_{it}</i>		-0.0401	0.0475	-0.0267	0.0524	-0.0398	-0.0400	-0.0397	-0.0157	-0.0479	-0.0240
		0.0553	0.0650	0.0595	0.0677	0.0553	0.0553	0.0553	0.0560	0.0556	0.0562
<i>dChair3_{it}*ln(PrivatefundingO)_{it}</i>			0.0015		0.0026						
			0.0040		0.0040						
<i>dChair4_{it}*ln(PrivatefundingO)_{it}</i>			-0.0122 **		-0.0118 **						
			0.0048		0.0048						
<i>dChair3_{it}*ln(NFPfundingO)_{it}</i>				-0.0078 **	-0.0080 **						
				0.0037	0.0037						
<i>dChair4_{it}*ln(NFPfundingO)_{it}</i>				-0.0031	-0.0019						
				0.0044	0.0044						
<i>dFemale_i*ln(PrivatefundingO)_{it}*dChair3_{it}</i>									-0.0012		0.0001
									0.0081		0.0082
<i>dFemale_i*ln(PrivatefundingO)_{it}*dChair4_{it}</i>									-0.0280 ***		-0.0311 ***
									0.0102		0.0103
<i>dFemale_i*ln(NFPfundingO)_{it}*dChair3_{it}</i>										-0.0087	-0.0091
										0.0065	0.0065
<i>dFemale_i*ln(NFPfundingO)_{it}*dChair4_{it}</i>										0.0120	0.0174 *
										0.0103	0.0104
<i>Constant term</i>	-0.0326	-0.1656	-0.2236	-0.2009	-0.2565	-0.1650	-0.1656	-0.1649	-0.1795	-0.1607	-0.1719
	0.2714	0.2711	0.2719	0.2715	0.2723	0.2712	0.2712	0.2712	0.2711	0.2712	0.2712
<i>Number of observations</i>	9097	9097	9097	9097	9097	9097	9097	9097	9097	9097	9097
<i>Number of scientists</i>	836	836	836	836	836	836	836	836	836	836	836
χ^2	2185.96	2231.62	2230.58	2237.66	2237.25	2231.39	2230.92	2230.54	2235.76	2234.8	2239.7
<i>sigma</i>	0.6921	0.6842	0.6844	0.6835	0.6836	0.6843	0.6844	0.6845	0.6840	0.6842	0.6843
<i>rho</i>	0.4798	0.4675	0.4677	0.4672	0.4672	0.4676	0.4678	0.4680	0.4675	0.4678	0.4682
<i>R2 within groups</i>	0.1385	0.1393	0.1392	0.1398	0.1398	0.1394	0.1394	0.1394	0.1399	0.1399	0.1406
<i>R2 overall</i>	0.3300	0.3409	0.3406	0.3411	0.3407	0.3411	0.3410	0.3411	0.3409	0.3408	0.3413
<i>R2 between groups</i>	0.4584	0.4730	0.4729	0.4728	0.4726	0.4731	0.4729	0.4731	0.4724	0.4722	0.4726

it is possible to argue that having a chair improve networking capability or funding amount of scientists.

In the second hypothesis we try to make a distinction between the effect of funding and having a research chair. By running regression model only on matched pairs of scientists, having a chair cannot be a proxy for criteria of matching (age, gender, and research field) anymore. We can verify the hypothesis 2 for industrial chair and research chairs appointed by research council but this hypothesis cannot be validated for 'Canada research chair' because its effect is still positive and significant even after matching. Some justification can be provided for this finding. The first is that Canada research chair intends to be prestige sign of research in Canada. Based on its mandate, the Canada research chair program aims to attract and retain some of most accomplished and promising minds in the world. It is more prestigious than other research chairs and other scientists may also have more willingness to conduct collaborative research with the Canada research chair holders. As the second justification, it should be noted that industrial chairs are appointed by firms to promote research, probably with major benefits for firms. In other words, this type of chair is not necessarily and originally designed for the sake of scientific publication. The chairs appointed by research councils may have quite similar characteristic. Looking at these chairs' description, most of chair holders are appointed as industrial chair. There are some evidence in literature indicating that industrial funding forces researchers to shift to more applied research, neglecting their normative responsibilities for knowledge development (Geuna & Nesta, 2003; Partha & David, 1994).

In addition to the effect of chair on scientific productivity, there are also some interesting results for other control variables in econometric model. Funding from different sources is always a significant determinant of scientific productivity, which has positive sign. Funding from private sector and funding from not-for-profit sector are directly put in regression equation while funding from public sector is first estimated by instrumental variables and then inserted to regression model.

The age of scientists seems to affect scientific productivity with an inverted-U shape pattern. However, considering its peak, which is 10 years old and less than the normal age for scientific activity, it is possible to argue that scientific productivity of scientists decreases in age. The gender of scientist, as another individual attribute, shows a significant impact. It indicates that women are less likely to publish journal paper compared with men. Both of these findings have some similar evidence in literature as discussed in previous section for age (Bonaccorsi & Daraio, 2003; Diamond, 1986; Levin & Stephan, 1991) and gender (Long, 1990).

The results verify the fixed effect of university and research discipline in addition to the year-specific effect on scientific production. Our regression analysis also tests the interactive effects of RHS variables. The first interactive effect is the interaction between gender and funding. From technical point of view, it is not possible to estimate the interactive effect with an endogenous variable in 2SLS models because its amount is estimated in the first stage and we are not using the raw value reported in dataset. However, we can estimate the effect of interaction with private funding and not-for-profit funding, which both are not significant. The only exception is in table 2 where the regression is run on whole dataset and interaction of gender and not-for-profit is negative and significant, which means that women may benefit from not-for-profit funding less efficiently compared with men.

The variables measuring interaction between having a chair and amount of funding are the next possible interaction in regression analysis, most of which are not significant. However, if there is significance, it is positive before matching and negative after matching. It refers to the more impact of funding for the chair people in general (complete data set) but when the chairs are compared to scientists, who are similar to them in terms of funding, gender, and research

field, they benefit from the funding less efficiently compared to non-chairs. The last group of interactive variables are the combination of two previous groups: interaction between funding, chair, and gender. There are some negative and significant effects for this type of interaction, showing the combined results of previous interactive variables.

Conclusion

In this article we show that having a research chair is a significant determinant of scientific publication when the regression is run over whole data set. As previously explained, a distinction should be made to clarify different attributes of research chair and their effect on scientific productivity. For instance, it is a fact that research chairs receive more grants due to their chair so the question here is to check if positive effect of research chair on scientific productivity remains significant after controlling for the funding amount of chair. To investigate the causality of this relationship, the matching technique is applied to control for some common characteristics of chair and non-chair scientists and to highlight the channel through which this positive effect has happened.

To conduct this matching technique, we only keep pairs of chair and non-chair scientists, matched together based on funding, gender, and research field, and delete the rest of scientists from data set. This methodology is effective to understand other attributes of research chair (except funding) that have significant and positive effect on scientific productivity. After such matching, the results show that the effect of Canada research chair on scientific productivity remains significant and positive while the effect of industrial chairs and the chairs appointed by Canada research council (NSER, SSHRC, and CIHR) become insignificant. This finding indicates that there are some special attributes in Canada research chair, which do not exist in other chairs. Those attributes may significantly push scientific productivity. Among other attributes, Canada research chairs may have better prestige to absorb talents or they are well designed to conduct scientific research for publication.

References

- Baird, L.L. (1986). What characterizes a productive research department? *Research in Higher Education*, 25(3), 211-225.
- Baird, L.L. (1991). Publication productivity in doctoral research departments: Interdisciplinary and intradisciplinary factors. *Research in Higher Education*, 32(3), 303-318.
- Beaudry, C. & Allaoui, S. (2012). Impact of public and private research funding on scientific production: The case of nanotechnology. [Working paper].
- Becker, G.S. (1962). Investment in human capital: a theoretical analysis. *The journal of political economy*, 70(5), 9-49.
- Bérubé, C. & Mohnen, P. (2009). Are firms that receive R&D subsidies more innovative? *Canadian Journal of Economics/Revue canadienne d'économique*, 42(1), 206-225.
- Blackburn, R.T., Behymer, C.E., & Hall, D.E. (1978). Research note: Correlates of faculty publications. *Sociology of Education*, 51(2), 132-141.
- Bonaccorsi, A., & Daraio, C. (2003). Age effects in scientific productivity. *Scientometrics*, 58(1), 49-90.
- Buchmueller, T.C., Dominitz, J., & Lee Hansen, W. (1999). Graduate training and the early career productivity of Ph.D. economists. *Economics of Education Review*, 18(1), 65-77. doi: 10.1016/s0272-7757(98)00019-3
- Cantu, F.J., Bustani, A., Molina, A., & Moreira, H. (2009). A knowledge-based development model: the research chair strategy. *Journal of Knowledge Management*, 13(1), 154-170.
- Courty, P., & Sim, J. (2012). What is the cost of retaining and attracting exceptional talents? Evidence from the Canada Research Chair program: Queen's Economics Department Working Paper.
- Crespi, G.A., & Geuna, A. (2008). An empirical study of scientific production: A cross country analysis, 1981–2002. *Research Policy*, 37(4), 565-579.
- Diamond, A.M. (1986). The life-cycle research productivity of mathematicians and scientists. *Journal of Gerontology*, 41(4), 520.
- Frey, B.S., & Rost, K. (2010). Do rankings reflect research quality? *Journal of Applied Economics*, 13(1), 1-38.
- Geuna, A. & Nesta, L. (2003). University patenting and its effects on academic research. *SEWPS Paper* (99).

- Gonzalez-Brambila, C., & Veloso, F.M. (2007). The determinants of research output and impact: A study of Mexican researchers. *Research Policy*, 36(7), 1035-1051.
- Henrekson, M., & Waldenström, D. (2007). How should research performance be measured: IFN Working Paper.
- Jordan, J.M., Meador, M., & Walters, S.J.K. (1988). Effects of department size and organization on the research productivity of academic economists. *Economics of Education Review*, 7(2), 251-255.
- Jordan, J.M., Meador, M., & Walters, S.J.K. (1989). Academic research productivity, department size and organization: Further results. *Economics of Education Review*, 8(4), 345-352.
- Katz, J. S., & Martin, B.R. (1997). What is research collaboration? *Research Policy*, 26(1), 1-18.
- Kyvik, S. (1995). Are big university departments better than small ones? *Higher Education*, 30(3), 295-304.
- Levin, S. G. & Stephan, P. E. (1991). Research productivity over the life cycle: evidence for academic scientists. *The American Economic Review*, 114-132.
- Long, J.S. (1990). The origins of sex differences in science. *Social Forces*, 68(4), 1297-1316.
- Long, J.S. (1992). Measures of sex differences in scientific productivity. *Social Forces*, 71(1), 159-178.
- Long, J.S., Allison, P.D., & McGinnis, R. (1979). Entrance into the academic career. *American Sociological Review*, 44(5), 816-830.
- Melin, G. & Persson, O. (1996). Studying research collaboration using co-authorships. *Scientometrics*, 36(3), 363-377.
- Newman, M.E.J. (2001a). Clustering and preferential attachment in growing networks. *Physical Review E*, 64(2), 025102.
- Newman, M.E.J. (2001b). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review-Series E*, 64(1; Part 2), 16132-16132.
- Niosi, J. (2002). Regional systems of innovation: Market pull and government push. *Holbrook, J.-A and D. Wolfe, (eds.) Knowledge, Clusters and Regional Innovation. Montréal & Kingston, McGill-Queen's University Press*, 39-55.
- Partha, D. & David, P.A. (1994). Toward a new economics of science. *Research Policy*, 23(5), 487-521.
- Pavitt, K. (2000). Why European Union funding of academic research should be increased: a radical proposal. *Science and Public Policy*, 27(6), 455-460.
- Pavitt, K. (2001). Public policies to support basic research: What can the rest of the world learn from US theory and practice? (And what they should not learn). *Industrial and corporate change*, 10(3), 761-779.
- Salter, A.J. & Martin, B.R. (2001). The economic benefits of publicly funded basic research: a critical review. *Research Policy*, 30(3), 509-532.
- Schimank, U. (2005). 'New Public Management' and the academic profession: Reflections on the German situation. *Minerva*, 43(4), 361-376.
- Turner, L. & Mairesse, J. (2005). Individual Productivity Differences in Public Research: How important are non-individual determinants? An Econometric Study of French Physicists' publications and citations (1986-1997). *Centre National de la Recherche Scientifique*.
- Van Raan, A.F.J. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62(1), 133-143.
- West, M.A., Smith, H., Feng, W.L., & Lawthom, R. (1998). Research excellence and departmental climate in British universities. *Journal of Occupational and Organizational Psychology*, 71(3), 261-281.
- Zhou, Y.B., Lü, L., & Li, M. (2012). Quantifying the influence of scientists and their publications: distinguishing between prestige and popularity. *New Journal of Physics*, 14, 033033.

Appendix 1 – Variable description.

<i>Variable name</i>	<i>Variable description</i>
<i>dChair1</i>	Dummy variables taking the value 1 if a scientist has a research chair awarded by industry (industrial chair)
<i>dChair2</i>	Dummy variables taking the value 1 if a scientist has a research chair awarded by Canadian funding agencies (NSERC, SSHRC, and CIHR)
<i>dChair3</i>	Dummy variables taking the value 1 if a scientist has a Canada research chair
<i>dChair4</i>	Dummy variables taking the value 1 if <i>dChair1</i> or <i>dChair2</i> are equal to 1
<i>dChair5</i>	Dummy variables taking the value 1 if any of <i>dChair1</i> , <i>dChair2</i> , or <i>dChair3</i> is equal to 1
<i>ln(PublicfundingO)</i>	Natural logarithm of the three-year average of public sector funding for the purpose of operational costs and direct expenditure of research
<i>ln(PrivatefundingO)</i>	Natural logarithm of the three-year average of private sector funding for the purpose of operational costs and direct expenditure of research
<i>ln(NFPfundingO)</i>	Natural logarithm of three-year average of funding from not-for-profit institutions (NFP) for the purpose of operational costs and direct expenditure of research
<i>ln(nb.Article)</i>	Natural logarithm of the yearly number of articles
<i>PubORank</i>	Normalized rank of scientist in the field in terms of three-year average of funding for the purpose of operational costs and direct expenditure of research
<i>PublRank</i>	Normalized rank of scientist in the field in terms of three-year average of articles count
<i>ln(totFund)</i>	Natural logarithm of three-year average of aggregate public sector funding in the field
<i>Age</i>	Age of a scientist
<i>dFemale</i>	Dummy variable taking the value 1 if the scientist is a woman and 0 otherwise
<i>dULaval, dUMcGill, ..., dUdeM</i>	Dummy variables indicating the university affiliation of researcher
<i>dMedical, dHumanities, ..., dScience</i>	Dummy variables indicating the field of researcher
<i>d2000, d2001, ..., d2012</i>	Dummy variables indicating the year

Drivers of Higher Education Institutions' Visibility: A Study of UK HEIs Social Media Use vs. Organizational Characteristics

Julie M. Birkholz¹, Marco Seeber¹ and Kim Holmberg²

{Julie.Birkholz, Marco.Seeber}@UGent.be

¹Centre for Higher Education Governance Ghent, Ghent University, Korte Meer 5, 9000, Gent (Belgium)

Kim.J.Holmberg@utu.fi

²Research Unit for the Sociology of Education, University of Turku, 20014 Turku (Finland)

Abstract

Social media is increasingly used in higher education settings by researchers, students and institutions. Whether it is researchers conversing with other researchers, or universities seeking to communicate to a wider audience, social media platforms serve as a tool for users to communicate and increase visibility. Scholarly communication in social media and investigations about social media metrics is of increasing interest for scientometric researchers, and to the emergence of altmetrics. Less understood is the role of organizational characteristics in garnering social media visibility, through for instance liking and following mechanisms. In this study we aim to contribute to the understanding of the effect of specific social media use by investigating higher education institutions' presence on Twitter. We investigate the possible connections between followers on Twitter and the use of Twitter and the organizational characteristics of the HEIs. We find that HEIs' social media visibility on Twitter are only partly explained by social media use and that organizational characteristics also play a role in garnering these followers. Although, there is an advantage in garnering followers for those first adopters of Twitter. These findings emphasize the importance of considering a range of factors to understand impact online for organizations and HEIs in particular.

Conference Topic

Science policy and research assessment, Country-level studies, Webometrics, Altmetrics

Introduction

The use of social media increases visibility of users (Constantinides & Zinck, 2011). This online visibility garners success and performance (Schindler & Bickar, 2005; Dellarocas, 2003; Duan et al., 2008). Less understood is the role of offline effects in garnering this visibility. For example, how do organizational characteristics influence an organization's visibility on social media? The understanding of the potential dual role of organizational characteristics and social media use in explaining visibility allows us to delineate how traditional characteristics such as status or reputation of organization play a role in generating attention on social media and how best to measure this impact.

We explore this through the lens of higher education. Social media is increasingly used in scholarly communication. Higher education institutions (HEIs), in particular, are increasingly using social media platforms as tools to communicate to prospective and current students, alumni and society at large (Gibbs, 2002; Helgesen 2008; Hemsley-Brown & Oplatka, 2006). Thus, the case of higher education and institutions' social media use in particular provides a valuable case to explore the possible dual role of organizational characteristics and the use of social media by these institutions in explaining garnered visibility.

In this paper we review literature on visibility of organizations and identify the potential role of social media use and organizational characteristics in explaining this visibility. We propose a number of hypotheses in which social media visibility is dependent on the two. To test these effects, we investigate 137 UK higher education institutions, collecting data of their Twitter activities and characteristics to explain social media visibility. Findings suggest that organizational characteristics of HEIs play a large role in their social media visibilities on

Twitter, compared to social media use alone. This emphasizes the importance of considering a range of factors in understand impact online for both organizations and HEIs in particular. This topic is of interest for scientometric researchers, as it is an additional avenue from bibliometrics to evaluate potential impact of a HEIs. In particular this work contributes to recent research on altmetrics. Altmetrics seeks to investigate the potential use of social media metrics for research evaluation and mapping of scholarly communication (Priem et al., 2010). The delineation of this mechanism advances our understanding of metrics validity and sheds light on the practical questions of how organizations can garner visibility online.

Social media and organizations

Organizational visibility is generated by the organization itself, and the users that engage with organizations. Organizational visibility is partly generated through word-of-mouth (WOM). WOM is the practice of communication where information is spread between individuals about a product or a service of a given organization (Richins, 1983). This mechanism allows individuals to share information and opinions to others about specific products, brands and services (Hawkins, Best, Coney, 2004; Westbrook, 1987) and to attach sentiment to messages. Positive WOM influences the awareness, image, decisions, evaluation and interest of potential consumers and stakeholders (Ozcan & Ramaswamy, 2004; Price, Feick & Guskey, 1995).

Organizations in particular are keen to attempt to achieve or maintain this positive WOM through different strategies of communication about the product or service they offer. With nearly half of all US internet users engaging on social networking sites (Smith, 2011), and with the numbers increasing worldwide, it is not a surprise that organizations are also getting involved in communicating via social media. The use of social media by organizations has largely been seen as marketing strategy to increase visibility (Constantinides & Zinck, 2011). Social media in particular serve as platforms for electronic WOM where entities spread and share information, but also as a medium where identification of organizational interests is transparent through online liking or following mechanisms (Dellarocas, 2003). Social media platforms serve as sites of social interaction, communication and marketing. This is achieved through socializing and networking online through text, images and videos. These platforms are largely made of user-generated content and facilitated through peer-to-peer communication and participation (Nambisan & Nambisan, 2008; Shankar & Malhotra, 2009).

A number of positive outcomes have been attributed to the use of social media by organizations. The use of social media platforms and thus consequent eWOM around a product or service of an organization influences attitudes, intentions and buying decisions (Schindler & Bickart, 2005; Goldsmith & Horowitz, 2006; Yao, Dresner & Palmer, 2009). The use of social media has also been attributed to increased economic impact (Chevalier & Mayzlin, 2006; Dellarocas, 2003). Recent work has questioned the impact of social media use on outcomes, suggesting that online content is solely a predictor of economic success, and not a factor that influences buying decisions (Chen et al., 2011; Duan et al., 2008). Follow-up studies suggest that user consult the Web for a confirmation of a decision they have made about a product, service or organization (Schindler & Bickart, 2005). Thus, this questions the explanatory power of social media use in garnering different outcomes, suggesting that other information about an organization or its product or service may play a role in understanding this garnered visibility online.

External to social media, the organization has a reputation, status and perceived legitimacy of an organization (Baum & Oliver, 1991; DiMaggio & Powell, 1983). Qualities such as status are said to determine a part of users'/consumers' expectations of future qualities of organizations (Podolny, 1993), which aid in defining the visibility and positions of an

organization in a field (Wry et al., 2009). Consequently, the degree of visibility of higher education institutions is not only dependent on the institution's use of social media for exposure, but also on certain organizational characteristics. Thus, we question: in addition to the use of social media platforms, how do organizational characteristics influence online visibility?

Higher education institutions and social media

In this paper we investigate organizations in the system of higher education. With higher education we mean the organizations that organize education and research, such as universities. Higher education is an industry in which consumers are often under informed in the sense that they cannot objectively evaluate the quality of the service before they actually “purchase it” (Jongbloed, 2003). Thus visibility about the organization is highly dependent on word-of-mouth practices to foster interest of potential students, research funding, and public support.

There is a rise of social media use by higher education institutions as tools in communicating information about the organization to prospective and current students, alumni and society at large (Gibbs, 2002; Helgesen, 2008; Hemsley-Brown & Oplatka, 2006). Social media fill a gap in the information that these groups cannot find in other forms of communication (Hemsley-Brown & Oplatka, 2006) such as alternative contact points for education and campus life (Yu et al., 2010; Mason & Rennie, 2007). Research shows that social media serves to fill a gap in the information that those interested in a university cannot find on the websites (Hemsley-Brown & Oplatka, 2006). Studies have found a significant relationship between those who logged onto the social media platform and the likelihood of them applying to the university (Hayes, Ruschman, & Walker, 2009). Thus, social media by higher education institutions serves said to play a positive role in garnering visibility through different methods.

On the other hand, recent studies in webometric studies of scholarly communication Web indicators or altmetrics have frequently been compared against more traditional indicators of research productivity (such as number of publications) and research impact (citations). Studies on the individual level found significant correlations between traditional bibliometric metrics, for instance research productivity and online visibility (Bar-Ilan, 2004; Thelwall & Tang, 2003). This relationship has been attributed to highly cited scholars producing more content on the web, which then attracted more attention (Thelwall & Harries, 2003). This has also been found in recent studies on HEIs, questioning how social media platforms play a lesser role than other forms of communication in attracting students in particular (Constantinides & Zinck Stagnoth, 2011), as well as the role of geographical proximity in the likelihood of universities in particular to link with other universities (Heimeriks & Van den Besselaar, 2006).

This is not necessarily striking given that HEIs have reputations external to the messages disseminated on social media platforms. Organizations are expected to capitalize on a baseline visibility as scholars have shown that organizations with a central position in the system, related to the organizational size, status and reputation, receive more attention from audiences and stakeholders (Wry et al., 2011, Podolny, 1993). Recent works in webometrics have also demonstrated that core organizational attributes matter in explaining online communication; where status, reputation and size are important predictors of hyperlink connections and centrality (Seeber et al., 2012, Lepori et al., 2013). Thus, using a social media platform does not alone garner visibility or interest from others. Given this we propose:

Hypothesis 1: Social media visibility can be explained by the social media use of the organization.

Hypothesis 2: The social media visibility of the organization can be explained by a HEIs organizational characteristics related to organizational size, status and reputation.

Hypothesis 3: The social media visibility of the organization can be explained by both the HEI's social media activity and organizational characteristics related to organizational size, status and reputation.

Methodology

We explore in this study UK universities, investigating both their Twitter activity and organizational characteristics. In selecting a social media platform where HEIs are active we have selected Twitter. Twitter is especially efficient for word-of-mouth marketing, given the ability to *re-tweet* – forward messages from users (Jansen et al., 2009) In addition tweets often contain expressions of sentiments (ibid), which makes it a valid source for identifying practices driven by potential eWOM. Following the theoretical framework, we assume that followers are a function of the organizational attributes and the social media use of the university.

Sample

Alike most European universities, UK universities are public institutions and the State and related funding bodies represent the most important funding sources.¹ On the other hand, UK universities are autonomous institutions, provided with strong decision making hierarchies and operating in a competitive system, they are expected to be able and in need of developing strategies to actively improve their position in the system (de Boer & Jongbloed, 2012; Seeber, et al., forthcoming). In turn, the UK Higher Education is a suitable case to explore what determines social media visibility in a quasi-market public system. Our sample includes 137 UK HEIs included in the European Micro Data dataset (Eumida) - a database containing the structural characteristics of 2,457 Higher Education institutions in twenty-eight European countries (Bonaccorsi et al., 2010; Eumida, 2009).²

Measures

We retrieved data from the HEIs' Twitter accounts manually. This data was collected on 24 November 2014 to measure the dependent variable of visibility and the independent variable - social media use. We also collected data on the organizational characteristics of the institutions, the second independent variable, for measuring a number of characteristics of the HEIs.

Visibility

We focus in this paper on social media visibility. This is a count variable that identifies the number of followers of each UK HEIs.

¹ HESA statistics on finance of UK universities available at: <https://www.hesa.ac.uk/>

² EUMIDA data refer to year 2007. Originally it included 148 universities, although four institutions have merged in the meanwhile, leading to a sample of 144. The Institute of Cancer Research and the London School of Hygiene and Tropical Medicine were excluded, as they are research institutes rather than HEIs; as well as the University of Southampton as it missed a value on coreness, one of the major predicting variables. Four outliers cases in terms of the number of followers were also excluded, leading to a sample of 137 UK HEIs; the University of Oxford, with 175,000 followers, The University of Cambridge 151,000, the Open university 100,000 and the London Business School 69,800, compared to a mean of 20,217 and standard deviation of 21,466.

Social media use

Scholarly communication in social media has been measured in a number of ways. Following literature suggesting a combination of activities we seek to identify attributes of the ways that HEIs use social media. Aguillo (2009) suggested using Web data as indicators related to 1) activity, 2) impact, and 3) usage. Indicators related to activity include measurements of the efforts made to actively create and establish a Web presence, while impact is the mentions on and linking from other websites. Usage is a proxy for the number of downloads or how users engage with the organization on the web. Given these metrics we sought to collect any query-able data on Twitter use. We collected data on the total number of tweets sent, the number of users that the HEIs themselves are following as a measure of their activity. Data was collected the date of HEI's first tweets obtained from the Twitter website³. In addition we collected data on the HEIs using Twitter to disseminate and share news and events or targeting students, as indicated by the HEIs in their profiles.

Organizational characteristics

We selected organizational characteristics that are deemed to be particularly relevant for the visibility of universities. We sought to identify on a number of measure of the universities' size, age, resources and status. The organizational features were constructed by using information from Eumida (Bonaccorsi, et al., 2010; Eumida, 2009). We considered, in particular; a) the size of the university, in terms of the number of staff units and undergraduate students; b) the university reputation in the core activities of research, measured through the scientific productivity and the research intensity, and teaching, measured through the teaching burden c) the university status, which is measured through the relational centrality of the university in the system (Owen-Smith & Powell, 2008). As control variables we considered; a) the discipline profile, as some disciplines may attract more attention than others because of the societal salience of the topics addressed, and b) the geographical context, in terms of the urban centrality of the city where the university is located, which may indirectly benefit the university's visibility. Table 1 describes the characteristics of each variable.

Results

Descriptive Statistics

Tables 2 and 3 present respectively the descriptive statistics and the Pearson correlation of the considered variables. The distribution of followers is moderately right skewed, as well as the number of Tweets, whereas the number of following is strongly right skewed. The days on Twitter is left skewed, as most universities started using twitter in early days and a small number of universities are late adopters (Table 2). Pearson correlations show that the number of followers is significantly correlated to most of the considered variables, and in particular to the status-coreness of the university (0.693), size measured by units of staff (0.642) and students (0.477), and scientific productivity (0.452). These organizational characteristics are strongly correlated with each other, so that high status universities are also large, and have a good scientific reputation. Variables of social media use are weakly correlated among each other and the organizational characteristics, with the highest correlations existing between the number of tweets and the size in terms of number of undergraduate students (0.264) (Table 3). The descriptive statistics show that the number of Twitter followers are characterized by over dispersion (i.e., the variance increases faster than the mean).

³ <https://discover.twitter.com/first-tweet#username>

Table 1. Description of organizational characteristics.

Size	The <i>number of total staff</i> (Full Time Equivalents measured in thousands), including academic as well as administrative and technical staff. The <i>number of undergraduate students</i> . (Eumida)
Reputation in research	Universities reputation in research activity is strongly related both to the <i>scientific productivity</i> , e.g. the quantity and quality of scientific publications. The indicator results from the product between the total number of publications multiplied by their field-normalized impact factor and divided by the number of academic staff. Data for two-thirds of the universities could be derived from the SCIMAGO institutional rankings for the year 2011 (http://www.scimagoir.com/), which is based on publications from the period 2005-2009; One-third of the universities are not covered since they had less than 100 publications in Scopus in the considered period. For these universities the indicator was set to zero. In fact, the scaling properties of research output (van Raan, 2007) maintain that the individual productivity tend to correlate with the organizational output, so that the indicator approaches zero when the level of output approaches the threshold of 100 publications. A second indicator of reputation in research considers the <i>research intensity</i> , as measured by the ratio between the number of PhD students over undergraduate students (Bonaccorsi, et al., 2007). (Eumida)
Reputation in teaching intensity	Teaching quality can be expected to be inversely related to the <i>teaching burden</i> , as measured by the ratio between the number of undergraduate students per unit of academic staff. (Eumida)
Status	University status is measured through the relational centrality or <i>coreness</i> in the system, estimated by considering web links connections between universities. Weblinks are receiving increasing attention in the study of inter-organizational relationships (Bar-Ilan 2009). European national higher education systems have been shown to conform to a core-periphery structure, where a status hierarchy is in place, core actors holding higher status and the <i>coreness</i> measuring the proximity to the network center (Borgatti & Everett, 1999; Lepori, et al., 2013; Owen-Smith & Powell, 2008).
Control: discipline profile	The disciplinary profile is defined by the share of academic staff employed in each of six subject domains considered in Science classification statistics (Eumida, 2009; Uoe, 2006). A Factor Analysis identifies three factors; separately employed as predicting variable. (Eumida)
Control: geographical context	The <i>Urban centrality</i> of the city where the university is located is measured through the Globalization and World Cities Network (GARC) scale of cities 2010 (Taylor, 2004) http://www.lboro.ac.uk/gawc/world2010.html . Accordingly, we ranked the universities with a numeric score from 9 (alpha++ cities) to 1 (gamma- cities), setting to zero the cities that are not in the list ^[1]

Table 2. Variables' descriptive statistics.

	<i>Mean</i>	<i>Median</i>	<i>Maximum</i>	<i>Minimum</i>	<i>Standard Deviation</i>
1 size - units of staff	2.001	1.665	9.498	68	1.675
2 size - undergraduate students	13.826	13.356	33.640	351	8.462
3 reputation - scientific productivity	274,66	72,50	1.828,00	0,00	389,03
4 reputation - research intensity	0,04	0,02	0,27	0,00	0,05
5 reputation - teaching burden	8,14	7,89	28,03	1,78	3,80
6 status - coreness	68	66	173	0	45
7 urban centrality	2,2	0,0	9,0	0,0	3,5
8 number of followers	17.189	15.900	46.200	1.233	10.085
9 number of tweets	6.792	5.598	19.000	300	4.220
10 days on twitter	1.918	2.019	2.644	305	342
11 number of following	1.312	832	12.700	107	1.506

Results of models

The dependent variable is represented by the number of Twitter followers, and assume that the number of followers is a function of the organizational attributes and the social media use of the university. Hence, we rely on techniques used for modelling count data for series of non-negative integers. If individual events are independent and their number is sufficiently large, the resulting probability distribution for the counts follows a Poisson distribution. Unlike linear regressions, the Poisson regression model does not assume that observations are normally distributed around the conditional mean, see Table 3. The descriptive statistics show that the number of Twitter followers are characterized by over dispersion (i.e., the variance increases faster than the mean). We then employ a negative binomial regression, which includes a parameter to model over dispersion. Table 4 presents the results of models: i) the empty model; ii) the model including the significant organizational characteristics; iii) the model employing the variables of social media use; iv) the full model including significant organizational characteristics and social media use variables.

Findings show that the social media are significant predictors of the number of followers, with the exception of the *number of following* (Hypothesis 1). In particular the *number of tweets* and *days on Twitter* have a positive effect; the orientation towards *news and events* has a positive and highly significant effect when compared to a *general* orientation. Findings also show that the organizational characteristics are predictive of the number of followers (Hypothesis 2). The *size in terms of undergraduate students* and the *research intensity* have a positive and strongly significant effect. Despite the lower correlation with the number of followers, these two measures are better predictors, respectively than the *size as number of staff units* and the *scientific reputation*. The variable on *status – coreness* is also strongly significant and positive. The *teaching burden*, the *discipline profile* as well as the *urban centrality* of the university' location are not significant predictors.

Comparatively speaking, the organizational characteristics model perform considerably better than the model on social media use.⁴ However, the final model (Hypothesis 3) displays that the better fit includes both organizational characteristics and social media use variables as regards the *number of tweets* and the *days on Twitter*.⁵ All variables have a positive and strongly significant effect. In order to assess the predictive capability of the full model we cannot rely on usual fit measures, like the R^2 , which assume a normal distribution. The model provides expected count values of followers, so that the fit can be judged by: a) computing a pseudo R^2 based on the formula: $1 - (\text{Total Sum Squared}/\text{Residual Sum Squared})$; b) computing the percentage of observed counts correctly predicted. The Pseudo R^2 is 0.66.⁶ Further, we consider the capability of the full model to correctly predict values below and above the median of 15,900 followers. The model correctly identifies 92% of the values below the median (*sensitivity*) and, when it predicts a value below the median, it is correct in 79% of the cases (*positive predictive value*). The performance is also good in terms of detecting the values above the median (67%, *specificity*); when the model predicts a value above the median, it is correct in 80% of the cases (*negative predictive value*). In sum, the overall predicting capability of the full model is fairly good. Figure 1. below displays a graphical depiction of these results, related to Twitter followers and organizational characteristics.

Binomial regression coefficients are exponential and multiplicative: if the coefficient for an antecedent is β , then the percentage change in the expected number of counts for unit a

⁴ Akaike Information Criterion - AIC (Akaike, 1998) of the null model is 2898.6, social media model AIC 2822.1 vs. organizational characteristics model AIC 2871.3, where lower values indicate a better fit.

⁵ Test for multicollinearity, VIF variance inflation factor, all variables well below the threshold of 10, the highest value observed for coreness at 2.62.

⁶ Pearson correlation between predicted and actual values is 0.826.

change in the antecedent is e^{β} . For instance, if the university “A” have 8,462 students more than university “B” (one standard deviation), it is predicted that A will have 1.16 times the number of followers of “B” (+16%).⁷ The observed coefficients confirm that both organizational characteristics and the specific use of social media have an important impact on the number of followers (Table 5).

Outliers

As a final test, we explore the capability of the full model to predict the four outlier cases that were excluded from the sample in a first stance. Whereas the number of followers of the Open University is reasonably well predicted (129,825 vs. 100,000 followers), the University of Oxford (60,180 vs. 175,000), the University of Cambridge (85,692 vs. 151,000), and the London Business School (12,624 vs. 69,800), attract a much larger number of followers than predicted by the model.

Table 3. Pearson correlation between the selected variables.

Table 3 - Pearson correlation between the selected variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 size - units of staff	1	,683**	,575**	,427**	-,291**	,804**	-,006	,513**	-,098	-,182*	,642**	,112	-,159	,183*
2 size - undergraduate students	,683**	1	,187*	-,065	,176*	,564**	-,152	,459**	,057	-,208*	,477**	,264**	,046	,106
3 reputation - scientific productivity	,575**	,187*	1	,495**	-,370**	,596**	,065	,465**	-,175*	-,100	,452**	-,035	-,107	,188*
4 reputation - research intensity	,427**	-,065	,495**	1	-,411**	,444**	,238**	,246**	-,038	-,019	,347**	-,185*	-,147	,029
5 reputation - teaching burden	-,291**	,176*	-,370**	-,411**	1	-,298**	-,107	-,173*	,095	-,056	-,230**	,090	,091	-,092
6 status - coreness	,804**	,564**	,596**	,444**	-,298**	1	-,046	,566**	,132	-,219*	,693**	,159	-,052	,145
7 urban centrality	-,006	-,152	,065	,238**	-,107	-,046	1	-,147	-,162	,044	-,052	-,290**	-,142	,017
8 discipline profile - factor 1	,513**	,459**	,465**	,246**	-,173*	,566**	-,147	1	,000	,000	,336**	,107	-,076	,085
9 discipline profile - factor 2	-,098	,057	-,175*	-,038	,095	,132	-,162	,000	1	,000	,066	,089	,060	-,069
10 discipline profile - factor 3	-,182*	-,208*	-,100	-,019	-,056	-,219*	,044	,000	,000	1	-,252**	-,121	-,114	-,058
11 number of followers	,642**	,477**	,452**	,347**	-,230**	,693**	-,052	,336**	,066	-,252**	1	,323**	,294**	,326**
12 number of tweets	,112	,264**	-,035	-,185*	,090	,159	-,290**	,107	,089	-,121	,323**	1	,120	,158
13 days on twitter	-,159	,046	-,107	-,147	,091	-,052	-,142	-,076	,060	-,114	,294**	,120	1	,033
14 number of following	,183*	,106	,188*	,029	-,092	,145	,017	,085	-,069	-,058	,326**	,158	,033	1

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 4. Negative Binomial regression models.

Table 4 - Negative Binomial regressions models

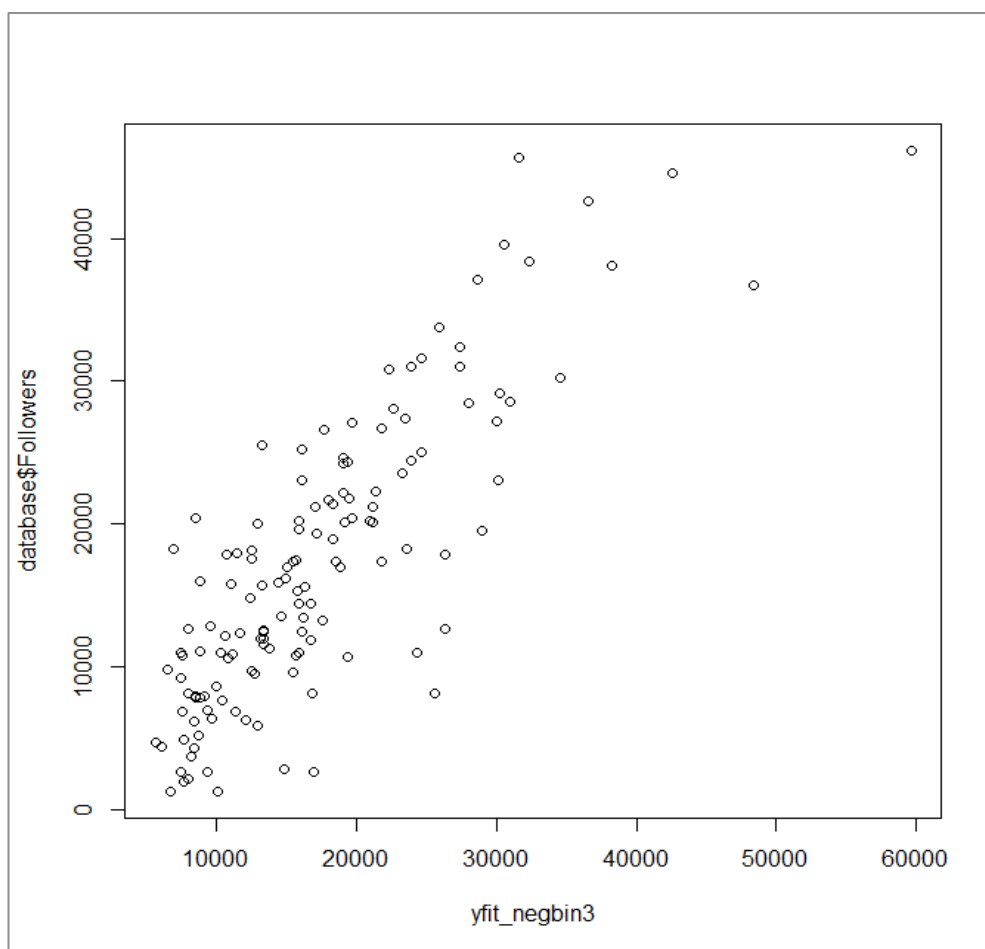
	Empty Model			Organizational variables Model			Social media use Model			Full Model		
	Estimate	S.E.	Pr(> z)	Estimate	S.E.	Pr(> z)	Estimate	S.E.	Pr(> z)	Estimate	S.E.	Pr(> z)
Intercept	9,752	0,054	<2e-16 ***	8,862	0,089	<2e-16 ***	8,371	289,300	<2e-16 ***	7,671	0,230	<2e-16 ***
size - undergraduate students				0,000023	0,000007	0,0007***				0,000018	0,000006	0,0035**
research intensity				2,774	1,088	0,01*				3,416	1,013	0,0007***
coreness				0,005	0,001	0,0002***				0,005	0,001	0,0004***
Tweets							0,00004	0,00001	0,0003***	0,00003	0,00001	0,0004***
days twitter							0,001	0,000	0,0003***	0,00053	0,00011	0,000002***
orientation: news and events							0,296	0,113	0,009**			
orientation: students							-0,312	0,183	0,09 .			
Null deviance	145,96	on 136 df		252,25	on 136 df		183,82	on 136 df		304,75	on 136 df	
Residual	142,49	on 136 df		142,25	on 133 df		144,15	on 132 df		141,37	on 131 df	
AIC:	2898,6			2822,1			2871,3			2798,4		
log-likelihood:	-2894,6			-2812,1			-2859,3			-2784,4		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

⁷ Changes in different antecedents have a multiplicative impact on expected number of followers. Hence, for instance, a university that is a standard deviation larger and research intensive than a university B will have 37% more followers ($1.16 \times 1.18 = 1.37$).

Table 5. Negative binomial regression model: comparing the impact of the variables.

		delta: standard deviation	proportion in number of followers
1	size - undergraduate students	8'462	1.16
2	research intensity	0.049	1.18
3	Status - coreness	45	1.24
4	Tweets	4'220	1.15
5	days twitter	342	1.20

**Figure 1. Results from full model.**

Discussion and conclusions

Findings indicate that both social media use and organizational characteristics explain the social media visibility of HEIs. Thus, organizations may be successful in garnering followers through their Twitter activity, but these high number of followers is also attributed to the organizational characteristics of size, status, and reputation. Notable is that these characteristics were better predictors of followers than the use of Twitter, suggesting that visibility is highly influenced by offline activities and traditional WOM, compared to eWOM. Although in regards to altmetrics – these online metrics do provide valid proxies for understanding dynamics, the addition of organization characteristics allows us to question how they serve as proxies, as the correlations suggest followers and following seem to be related to organizations size, and reputation, although the organizations own activities of tweeting and experience on Twitter are not related. That does not discard the power of social

media platforms as a tool for garnering visibility, although emphasizes that it is not a replacement for building reputation external to online domains.

Findings show that the social media are significant predictive of the number of followers, with the exception of the *number of following* (Hypothesis 1). In particular the *number of tweets* and *days on twitter* have a positive effect. Findings also show that the combined organizational characteristics are predictive of the number of followers (Hypothesis 2). The *size in terms of undergraduate students* and the *research intensity* have a positive and strongly significant effect. Despite the lower correlation with the number of followers, these two measures are better predictors, respectively than the *size as number of staff units* and the *scientific reputation*. The variable on *status – coreness* is also strongly significant and positive. The *teaching burden*, the *discipline profile* as well as the *urban centrality* of the university's location are not significant predictors.

In addition to the specific a number of notable findings emerged with regards to the specific variables. First, the importance of length of time on Twitter suggests a “first mover advantage”, where first adopters have yielded higher numbers of followers. HEIs Twitter accounts that had an orientation towards news and events play a more significant role in garnering online visibility through followers. Secondly, in regards to the organizational characteristics *size in terms of undergraduate students* and *research intensity* played the most significant role in explaining online followers. These two measures reflect the two core tasks of HEIs – research and education. That is HEIs that are able to attract a high number of students as well as sustain a higher number of PhD candidate to conduct research, which again garners increased social media visibility.

This study provides clear support for a causal mechanism that stipulates that both organizational characteristics and social media use explain social media visibility as measure by followers. This provides additional evidence to scientometricians of the importance of considering a combination of metrics in explaining impact and scholar impact in particular. Although, in this research we have analyzed basic descriptors. There is margin for improving explanation of social media use. Future research should investigate, for instance, the content of tweets, as well as the strategies for managing eWOM (Bao & Chang 2014). In addition, the existence of a few outliers suggests that few actors attract a disproportionally high attention from the public. Future research may investigate why this occurs. Given the state of literature we did not have evidence at the onset of our model to suggest an interaction effect, although given that the explanatory power of an organizations social media visibility is explained by both organizational characteristics and social media use, an interaction effect is a natural next step. For example, to investigate the effect of social media use by HEI on (social media) visibility is enhanced in HEIs with a large size, high status and high reputation.

References

- Aguillo, I. (2009). Measuring the institutions' footprints in the web. *Library High Tech*, 27(4), 540-556.
- Akaike H. (1998). Information theory and an extension of the maximum likelihood principle, in (eds) *Selected Papers of Hirotugu Akaike*, 199-213. New York: Springer.
- Bao, T., & Chang, T. L. S. (2014). Finding disseminators via electronic word of mouth message for effective marketing communications. *Decision Support Systems*, 67, 21-29.
- Baum, J. A. C., & Oliver, C. (1991). Institutional linkages and organizational mortality. *Administrative Science Quarterly*, 36, 187-218.
- Bar-Ilan, J. (2004). A microscopic link analysis of academic institutions within a country – the case of Israel. *Scientometrics*, 59(3), 391-403.
- Bar-Ilan J. (2009). Infometrics at the beginning of the 21st century - A review. *Journal of Infometrics*, 2(1), 1-52.
- Bonaccorsi A., Daraio, C., Lepori, B. & Slipersaeter, S. (2007). Indicators on individual higher education institutions: addressing data problems and comparability issues. *Research Evaluation*, 16(2), 66-78.

- Bonaccorsi A., Lepori, B., Brandt, T., De Filippo, D., Niederl, A., Schmoch, U., Schubert, T. & Slipersaeter, S. (2010). Mapping the European higher education landscape. New insights from the EUMIDA project. Science and Technology Indicators Conference, Leiden, the Netherlands, 9-11 September.
- Borgatti S. P. & Everett, M. G. (1999). Models of core/periphery structures. *Social Networks*, 21, 375-395.
- Chevalier, J. A. & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, 43(3), 345-354.
- de Boer H. & Jongbloed, B. (2012). A Cross-National Comparison of Higher Education Markets in Western Europe, in A. Curaj, P. Scott, L. Vlasceanu and L. Wilson (eds) *European Higher Education at the Crossroads: Between the Bologna Process and National Reforms*, 553-571. Dordrecht: Springer Netherlands.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48, 147-160.
- Chen, Y., Fay, S., & Wang, Q. (2011). The role of marketing in social media: How online consumer reviews evolve. *Journal of Interactive Marketing*, 25(2), 85-94.
- Constantinides, E., & Zinck Stagno, M. C. (2011). Potential of the social media as instruments of higher education marketing: a segmentation study. *Journal of Marketing for Higher Education*, 21(1), 7-24.
- Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, 49(10), 1407-1424.
- Duan, W., Gu, B., & Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *Journal of Retailing*, 84(2), 233-242.
- Eumida (2009). EUMIDA Handbook - Collection of institutional-level data on tertiary educational institutions at the European level. Eumida.
- Gibbs, G. (2002). Institutional strategies for linking research and teaching. *Exchange*, 3, 8–11.
- Goldsmith, R. E., & Horowitz, D. (2006). Measuring motivations for online opinion seeking. *Journal of Interactive Advertising*, 6(2), 2-14.
- Hawkins, D.I., Best, R. & Coney, K.A. (2004). *Consumer Behavior: Building Marketing Strategy*. 9th ed. Boston: McGraw Hill.
- Hayes, T. J., Ruschman, D., & Walker, M. M. (2009). Social networking as an admission tool: A case study in success. *Journal of Marketing for Higher Education*, 19(2), 109-124.
- Heimeriks, G. & Van den Besselaar, P. (2006). Analyzing hyperlinks networks: the meaning of hyperlink based indicators of knowledge production. *Cybermetrics*, 10(1), paper 1.
- Helgesen, Ø. (2008). Marketing for higher education: A relationship marketing approach. *Journal of Marketing for Higher Education*, 18(1), 50-78.
- Hemsley-Brown, J., & Oplatka, I. (2006). Universities in a competitive global marketplace: systematic review of the literature on higher education marketing. *International Journal of Public Sector Management*, 19(4), 316-338.
- Jansen, B.J., Zhang, M., Sobel, K. & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, 60(11), 2169-2188.
- Jongbloed, B. (2003). Marketisation in higher education, Clark's triangle and the essential ingredients of markets. *Higher Education Quarterly*, 57(2), 110-135.
- Lepori B., Barberio, V., Seeber, M. & Aguillo, I. (2013). Core-periphery structures in national higher education systems. A cross-country analysis using interlinking data, *Journal of Infometrics*, 7(3), 622-34.
- Mason, R. & Rennie, F. (2007). Using web 2.0 for learning in the community. *Internet and Higher Education*, 10, 196–203.
- Nambisan, S., & Nambisan, P. (2008). How to Profit From a Better Virtual Customer Environment. *MIT Sloan Management Review*, 49(3), 53.
- Owen-Smith J. & Powell, W. W. (2008) 'Networks and Institutions', in R. Greenwood, C. Oliver, K. Shalin & R. Suddaby (eds) *The Sage handbook of organizational institutionalism*, 594-621. London.
- Ozcan, K. & Ramaswamy, V. (2004) Word-of-mouth as dialogic discourse: A critical review, synthesis, new perspective, and research agenda. Working Paper. Retrieved June 15, 2015 from: <http://kerimcanozcan.com/portal/downloads/Word-ofMouth%20as%20Dialogic%20Discourse.pdf>.
- Price, L. L., Feick, L. F., & Guskey, A. (1995). Everyday market helping behavior. *Journal of Public Policy & Marketing*, 14(2), 255-266.
- Priem J., Taraborelli, D., Groth, P., & Neylon, C. (2010). alt-metrics: a manifesto. Retrieved June 15, 2015 from: <http://altmetrics.org/manifesto/>.
- Podolny, J. M. (1993). A status-based model of market competition. *American Journal of Sociology*, 98(4), 829-872.
- Richins, M. L. (1983). Negative word-of-mouth by dissatisfied consumers: a pilot study. *The Journal of Marketing*, 47(1), 68-78.

- Schindler, R. M., & Bickart, B. (2005). Published word of mouth: Referable, consumer generated information on the Internet. In *Online consumer psychology: Understanding and influencing consumer behavior in the virtual world*, 35-61.
- Seeber M., Lepori, B., Montauti, M., Enders, J., De Boer, H., Weyer, E., Bleiklie, I., Hope, K., Michelsen, S., Nyhagen, G., Frohlich, N., Scordato, L., Stensaker, B., Waagene, E., Dragsic, Z., Kretek, P., Krücken, G., Magalhães, A., Ribeiro, F., Sousa, S., Veiga, A., Santiago, P., Marini, G. & Reale, E. (2015). European Universities as Complete Organizations? Understanding Identity, Hierarchy and Rationality in Public Organizations. *Public Management Review*.
- Seeber, M., Lepori, B., Lomi, A., Aguillo, I., & Barberio, V. (2012). Factors affecting web links between European higher education institutions. *Journal of Informetrics*, 6(3), 435-447.
- Shankar, V., & Batra, R. (2009). The growing influence of online marketing communications. *Journal of Interactive Marketing*, 23(4), 285-287.
- Smith, A. (2011). Why Americans use social media. Pew Research Internet Project. Retrieved June 15, 2015 from: <http://www.pewinternet.org/2011/11/15/why-americans-use-social-media/>.
- Uoe (2006). *UOE data collection on education systems. Manual: Concepts, definitions, classifications*, Montreal, Paris, Luxembourg: UNESCO, OECD, Eurostat.
- Taylor P. J. (2004). *World City Network: a Global Urban Analysis*. London: Routledge.
- Thelwall, M. & Harries, G. (2003). Do the Web sites of higher rated scholars have significantly more online impact? *Journal of the American Society for Information Science and Technology*, 55(2), 149-159.
- Thelwall, M. & Tang, R. (2003). Disciplinary and linguistic considerations for academic Web linking: An exploratory hyperlink mediated study with Mainland China and Taiwan. *Scientometrics*, 58(1), 155-181.
- van Raan A.F.J. (2007). Bibliometric statistical properties of the 100 largest European universities: prevalent scaling rules in the science system. Available: *arXiv:0704.0889*.
- Westbrook, R. A. (1987). Product/consumption-based affective responses and post purchase processes. *Journal of Marketing Research*, 24(3), 258-270.
- Wry, T., Lounsbury, M., & Glynn, M.A. Legitimizing new categories of organizations: Stories as distributed cultural entrepreneurship. *Organization Science*, 22, 449-463.
- Yu, A., Tian, S., Vogel, D. & Kwok, R (2010). Embedded social learning in online social networking, in ICIS 2010 Proceedings 2010, Retrieved June 15, 2015 from: http://aisel.aisnet.org/icis2010_submissions/100.
- Yao, Y., Dresner, M., & Palmer, J. W. (2009). Impact of Boundary-Spanning Information Technology and position in Chain on Firm Performance. *Journal of Supply Chain Management*, 45(4), 3-16.

A Computing Environment to Support Repeatable Scientific Big Data Experimentation of World-Wide Scientific Literature

Bob G. Schlicher¹, James J. Kulesz², Robert K. Abercrombie³, and Kara L. Kruse⁴

¹schlicherbg@ornl.gov, ²jim.kulesz@gmail.com, ³abercrombie@ornl.gov, ⁴krusekl@ornl.gov

Oak Ridge National Laboratory, Computational Sciences and Engineering Division, 1 Bethel Valley Road, Oak Ridge, TN 37830-6085 (USA)

Abstract

A principal tenet of the scientific method is that experiments must be repeatable. This tenet relies on *ceteris paribus* (i.e., all other things being equal). As a scientific community, involved in data sciences, we must investigate ways to establish an environment where experiments can be repeated. We can no longer allude to where the data comes from, we must add rigor to the data collection and management process from which our analysis is conducted. This paper describes a computing environment to support repeatable scientific big data experimentation of world-wide scientific literature, and recommends a system that is housed at the Oak Ridge National Laboratory in order to provide value to investigators from government agencies, academic institutions, and industry entities. The described computing environment also adheres to the recently instituted digital data management plan, which involves all stages of the digital data life cycle including capture, analysis, sharing, and preservation, as mandated by multiple United States government agencies. It particularly focuses on the sharing and preservation of digital research data. The details of this computing environment are explained within the context of cloud services by the three layer classification of “Software as a Service”, “Platform as a Service”, and “Infrastructure as a Service”.

Conference Topic

Science policy and research assessment, Methods and techniques

Introduction¹

The scientific policy and research assessment community is investigating methods and techniques to establish an environment where experiments can be repeated through the use of data management. This approach attempts to ensure the integrity of scientific findings and the processes from which scientific literature analysis is conducted.

Data Science is the study of the generalizable extraction of knowledge from data (Dhar, 2013). From this definition, scientific development thus becomes the piecemeal process by which these items have been added, singly and in combination, to the ever growing stockpile that constitutes scientific technique and knowledge (Kuhn, 1970). Scientific literature analysis, or Scientometrics, is the study of measuring and analysing science, technology and innovation. Organizations, such as Thomson Reuters, have long used these analyses to identify the most influential papers or researchers in a field. Recently, Foresight and Understanding from Scientific Exposition (Murdick, 2011) takes this further by mining millions of papers and patents in both English and Chinese, two of the most commonly used languages in scientific literature (Readron, 2014).

¹ This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the United States Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

Scientometrics and its related research activities in today's world make extensive use of digital research data. The data management of this digital research data is, in essence, the quintessential requirement for repeatable scientific experimentation. This term, digital research data, encompasses a wide variety of information stored in digital form including: experimental, observational, and simulation data, codes, software and algorithms, text, numeric information, images, video, audio, and associated metadata. It also encompasses information in a variety of different forms including raw, processed, and analysed data, and published and archived data ("Statement on Digital Data Management," 2014). More specifically, research data are defined in regulation ("Intangible property - Code of Federal Regulations 2 CFR 200.315," 2014), continuing the definition in further statutes and United States Government Directives ("2 CFR 215 - Uniform Administration Requirements for Grants and Agreements With Institutions of Higher Education, Hospitals, and Other Non-Profit Organizations (OMB Circular A-110) ", 2012) as follows:

- “Research data is defined as the recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues. This 'recorded' material excludes physical objects (e.g., laboratory samples). Research data also do not include:
 - Trade secrets, commercial information, materials necessary to be held confidential by a researcher until they are published, or similar information which is protected under law; and
 - Personnel and medical information and similar information, which the disclosure would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study.”

Purpose of the Study

When addressing the reality of allocating the scarce resources of the current research budget constraints, the current institutions of science today operate, essentially the same, as from the time period just after the Second World War (Azoulay, 2012). Azoulay further argues it would be a fortuitous coincidence if the systems that served us so well in the twentieth century were equally adapted to twenty-first-century needs. Such is not the case. To leverage these finite resources and to adhere to the principle of the scientific method that all experiments must be repeatable, we, as a scientific community must investigate ways to establish environments where experiments can be repeated. We can no longer allude to from where the data come, we must add rigor to the data collection and data management process from which our analysis is conducted.

Data management involves all stages of the digital data life cycle including capture, analysis, sharing, and preservation. The focus of this statement is the sharing and preservation of digital research data. The following principles apply to the effective management of digital research data ("Statement on Digital Data Management," 2014):

- Effective data management has the potential to increase the pace of scientific discovery and promote more efficient and effective use of government funding and resources. Data management planning should be an integral part of research planning.
- Sharing and preserving data are central to protecting the integrity of science by facilitating validation of results and to advancing science by broadening the value of research data to disciplines other than the originating one and to society at large. To the greatest extent and with the fewest constraints possible, and consistent with the requirements and other principles of this statement, data sharing should make digital

research data available to and useful for the scientific community, industry, and the public.

- Not all data need to be shared or preserved. The costs and benefits of doing so should be considered in data management planning.

Procedure for a Computing Environment to Support Repeatable Scientific Big Data Experimentation

A data management plan is a formal document that outlines how a research institution and program will handle data both during research and after the project is completed ("Data management plan," 2014). The goal of a data management plan is to consider the many aspects of data management, metadata generation, data preservation, and analysis before the project begins. This ensures that data are well-managed in the present and prepared for preservation in the future. Multiple United States government agencies now require proposals submitted to include a supplementary document labelled "Data Management Plan" (Collins, 2014; "Dissemination and Sharing of Research Results," 2010). These supplementary documents describe how the proposal will conform to scientific policy on the dissemination and sharing of research results.

FUSEnet is a data analytics cloud specializing in managing both data and computational processes for assessing technical knowledge for identifying emergent technologies and capabilities. Under a multi-year United States Government research effort sponsored by Intelligence Advanced Research Projects Activity (IARPA), the overall goal of the FUSE program is to produce a new capability to accelerate the process of identifying and prioritizing emerging technologies across the globe (Murdick, 2011). The FUSE Program was established to develop automated methods that aid in the systematic, continuous, and comprehensive assessment of technical emergence using information found in published scientific, technical, and patent literature. A concise description is as follows (Murdick, 2011):

A fundamental hypothesis of the FUSE Program is that real-world processes of technical emergence leave discernible traces in the public scientific, technical, and patent literature. FUSE envisions a system that can (1) process the massive, multi-discipline, growing, noisy, and multilingual body of full-text scientific, technical, and patent literature from around the world; (2) automatically generate and prioritize technical terms within emerging technical areas, nominate those that exhibit technical emergence, and provide compelling evidence for the emergence; and (3) provide this capability for literature in English and at least two non-English languages. Technology developed from the FUSE Program would automatically nominate both known and novel technical areas based on quantified indicators of technical emergence with sufficient supporting evidence and arguments for that nomination. The FUSE Program also addresses the vital challenge of validating such a system, using real world data.

FUSEnet is currently a government system hosted by ORNL that stores unclassified, copyright-protected scientific information and provides remote access for approved users to analyse the stored data within a cloud computing environment to satisfy the research objectives of the IARPA FUSE Program. A key tenet within FUSEnet is that data integrity and availability is maintained. An ORNL developed "data diode" embedded within FUSEnet gateways allows access to protected data, but prevents data removal by users. As necessary, a mechanism for approved data export is built into the system architecture. Also by design, the activities and work products of individual user teams are segregated from each other in the cloud computing virtual environment.

FUSEnet Capabilities

The FUSEnet computing environment is based on the Cloud service model. These models are usually described by a three layer classification called SPI for SaaS, PaaS, and IaaS (Tian & Zhao, 2015) and adapted as follows:

- SaaS – Software as a Service: applications that are available on-demand.
- PaaS – Platform as a Service: refers to a computing platform of software components and middleware that are used by end-users to develop and manage their cloud applications. Typically, cloud providers at this layer offer databases, web servers, development environments, and application monitoring tools.
- IaaS – Infrastructure as a Service: physical or virtual machines with access to data storage and other operating system services. The cloud user is typically expected to install and maintain operating-system images.

The unique processing capabilities of FUSEnet are in the SaaS and PaaS levels. The IaaS capabilities were established with off-the-shelf software and hardware solutions as a result of understanding the operational needs of FUSEnet users, big data analytics, and optimizing central processing unit (CPU) and input/output (I/O) performance. One of the major challenges with the computing environment is with moving large volumes of data (terabytes) to and from the disk storage to the CPUs for processing. This challenge is met with ever increasing improvements and replacements for the IaaS without having any operating impact on the SaaS or PaaS layers. FUSEnet demonstrated this with an improvement in the data I/O transfer by replacing the disk storage system over its earlier version. Further, FUSEnet SaaS and PaaS software can be hosted on commercial IaaS platforms that meet the requirements for its intended usage.

A summary of the FUSEnet benefits and capabilities that support repeatability of big data experiments includes:

- An organized repository of 100 million published scientific and patent documents,
- Technical in-house expertise for maintenance of data pertaining to integrity and availability, pedigree, and version control,
- Reliable data sources including data provided by, Thomson Reuters, Lexis-Nexis, Elsevier, Institute of Electrical and Electronics Engineers (IEEE), Nature Publishing Group, PubMed Central, and others,
- Technical expertise with the format and details of the data, and
- Four analytical software applications with evidentiary traceability and indicators for assessing repeatability:
 - Assess and forecast technical research and technology developments,
 - Reverse-search the events contributing to a technology or development,
 - Drill down the evidence supporting the assessment and forecast,
 - Remote end-user workspaces ready-to-run the applications and the analytics platform,
 - Multiple analytics capabilities including Natural Language Processing (NLP), Parts-of-Speech (PoS) detectors, deduplication, belief network modelling, and machine learning,
 - Operation of the system with 24/7 and 99.8% availability within domain-specific expertise with the current ORNL technical staff,
 - Rapid custom development to meet unique end-user analytics requirements, and
 - Immediate data protection for the repository and custom end-user data.

The FUSEnet SaaS Level

At the SaaS level, four unique software applications perform automated technical assessments for supporting the detection and forecasting. Each of these applications process and analyse published scientific and engineering papers that are made available in the FUSEnet data repository. Unlike previous approaches to detecting emergence, which are based on the citation analysis of papers and patents (Bettencourt et al., 2008; Huang et al., 2014), the following application systems extract information from the text of publications and patents, identifying authors, their affiliations, addresses, as well as classifying types of organizations and publications. Although these applications have the same objectives, their analytical techniques are uniquely different and hence provide different insights into the organization and search of the data (Babko-Malaya et al., 2013). These analysis techniques include: feature extraction (Michaelis et al., 2012), time series analysis, sentiment and network analysis (Fürstenau & Rambow, 2012), and emergent detection and prediction (Brock et al., 2012), among others. The four main applications developed within the FUSEnet system are ARBITER from BAE Systems, Copernicus from SRI International, Emerge from BBN, and DETAIL from Columbia University.

The FUSEnet PaaS Level

The aforementioned SaaS applications use a variety of tools and libraries at the PaaS level. While the SaaS level in FUSEnet is the automated assessment, the FUSEnet PaaS computing platform can best be described as a “Network Analysis” (Otto & Rousseau, 2002) and text analytics platform. Text analysis uses statistical pattern learning to find patterns and trends from text data (in our case, scientific literature and patents). A summary of several key tools that FUSEnet provides are in Table 1. A selection of software libraries for network analysis and text analysis in FUSEnet, available for ensuring that experiments can be repeated, is shown in Table 2.

The FUSE Program licensed and installed a large number of scientific papers and patents from several suppliers in multiple languages including English and Chinese. The data sets include bibliographic citations of journal articles (108+ million), full text journal articles (5+ million), patent backfile records (14+ million at beginning of 2013 for the US and China), and updates to the patent backfile records, (51+ million for the US and China). A backfile is a single file containing the original patent application data plus all updates to the patent (both by the originator and by the patent office) up to the time the backfile was created.

Fig. 1 illustrates the large increase in scientific journal articles and patent applications as included in the FUSE research system during the past two decades. The number of Chinese patent applications is increasing dramatically and has now surpassed the number of US patent applications. Also, the number of Chinese journal articles is increasing at a rate faster than the rest of the world.

Table 1. FUSEnet PaaS support software packages.

	<i>FUSEnet PaaS Analytics Tool</i>	<i>Technical Usage</i>	<i>SaaS application that uses it</i>
1	MySQL ²	SQL ³ database typically used to store document, term, and author data.	Emerge, ARBITER
2	MongoDB ⁴	Document-oriented, NoSQL database used to store extracted entities and indicator-specific data.	Emerge, Copernicus
3	MALLET	Machine Learning and NLP ⁵ Toolkit for Java. Provides topic modelling for document clustering.	Emerge
4	Sofia-ml	Fast incremental machine learning algorithm. Provides clustering of documents from topic models generated by MALLET.	Emerge
5	Lucene IR system	Used for its indexing engine.	Emerge
6	Scikit-learn	Machine learning models.	Emerge
7	Tomcat/Solr Web Server	Used for Term indexing.	ARBITER
8	Apache ActiveMQ ⁶	Messaging and integration patterns.	ARBITER
9	Cassandra	NoSQL database.	ARBITER
10	Virtuoso	RDF ⁷ triple storage.	ARBITER
11	OpenRDF/Sesame	RDF processing including parsing, storing, reasoning and querying.	ARBITER
12	Spring Framework	Used for Integration using JMS.	ARBITER
13	Lucene/Solr	Document level information search, retrieval and storage engine.	ARBITER, DETAiLS
14	Open NLP	Machine learning based toolkit for processing natural language text.	ARBITER
15	Netica	Used for working with belief networks and influence diagrams.	ARBITER
16	Elasticsearch	Extension on Lucene that provides search and analytics.	Copernicus
17	Hadoop 2+	Used for extract, transform, and load (ETL) and de-duplication processing.	Copernicus
18	Berkeley Parser	Sorts and assigns words in sentences into subjects, verbs, and objects.	DETAiLS
19	Duke	Deduplication engine written in Java operating with Lucene.	DETAiLS
20	Stanford Chinese Word Segmenter	Split Chinese text into a sequence of words.	DETAiLS
21	Stanford Part-of-Speech (POS) Tagger	Reads text and assigns parts of speech to each word (noun, verb, adjective, etc.).	DETAiLS
22	UIMA	Unstructured Information Management Architecture (UIMA) is a general framework for analysis of unstructured information and its integration with search technologies.	DETAiLS
23	Weka	Machine learning software written in Java for data analysis and predictive modelling.	DETAiLS

² MySQL is a well-known relational database manager used in a wide variety of systems, including Twitter, Wikipedia, Facebook, Google, Wordpress, and countless more websites and other applications.

³ SQL (Structured Query Language) is a special-purpose programming language designed for managing data held in a relational database management system (RDBMS),

⁴ MongoDB is a document-oriented, NoSQL database.

⁵ NLP is Natural Language Processing where algorithms are used to derive meaning from human language.

⁶ Apache ActiveMQ is an open source message broker written in Java together with a full Java Message Service (JMS) client.

⁷ RDF is Resource Description Framework and is used to express data in subject-predicate-object expressions.

Table 2. Subset of FUSEnet software libraries for social network and text analysis.

	<i>Library/Package</i>	<i>Description</i>	<i>SaaS application that uses it</i>
1	Arpack	Linear algebra routines for Java	Emerge
2	JDOM	XML processing library for Java	Emerge
3	Jwnl	Java WordNet library	Emerge, ARBITER
4	Matrix-toolkits-java	Linear algebra data structures for Java	Emerge
5	BLAS	Linear algebra subroutines	Emerge
6	LAPACK	Linear algebra data structures and subroutines	Emerge
7	Libquadmath	High-precision math libraries	Emerge
8	Beanshell	Scripting for Java	Emerge
9	Trove4j	High-performance data structures for Java	Emerge
10	JGraphT	Graphical data structures and algorithms for Java	Emerge
11	JUNG	Java Universal Network/Graph Framework	ARBITER
12	R	Development environment for statistical computing and graphics	ARBITER

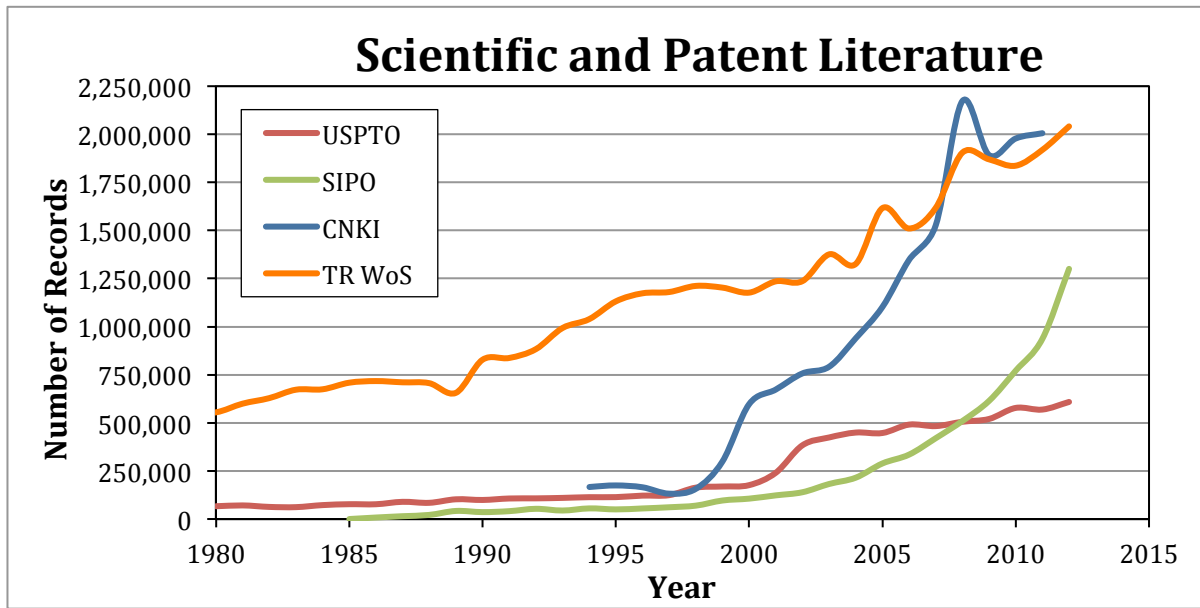


Figure 1. Number of records per year for the four largest datasets in the FUSEnet collection including patent records from the US (USPTO) and Chinese (SIPO) patent offices (i.e. number of backfile records at the beginning of 2013) and journal article citations from China (CNKI) and Thomson Reuters' Web of Science (TR WoS).

The FUSEnet IaaS Level

The deployed second generation FUSEnet at ORNL has the following summary specifications:

- 770 gigaFLOPS⁸ of maximum performance,
- 16 blade servers (plus 2 support blades), each with 2 CPUs, each with 6 cores, totalling 192 cores, or processors; additional blade with USB 3.0 for dedicated data transfer/export,
- 3.07 TB of RAM w/ 192 GB per node,
- Disks:
 - EMC Isilon: 340 TB (useable; includes 6.4 TB SSD) running NFS over 10 Gb/s Ethernet,
 - HP LeftHand: 260 TB of effective disk storage; will be reconfigured for backup and
 - Isilon disk I/O up to 1 gigabyte/sec per blade,
- Networking: Flex-10 modules totalling 160 Gbits/sec bandwidth per enclosure x 2 enclosures (theoretical maximum),
- Virtualized computing space through VMware⁹,
- Access and control policies enforced by ORNL Computing Data Center, and
- Call Center and metrics for service quality.

Table 3. Characteristics of cloud providers and applicability to FUSEnet requirements.

	<i>Vendors</i>	<i>Cloud Offering Overview</i>	<i>Applicability to FUSEnet</i>
1	Amazon Web Services	Overall market leader offering virtual servers, MapReduce (Hadoop) for search engine, large data storage, SQL databases, NoSQL databases, mobile integration, business applications including email, payment systems, and workflow.	PaaS (databases), IaaS
2	Google Cloud Platform	App Engine web application platform (PaaS), virtual machines, file storage, SQL databases, NoSQL, big dataset support, mobile integration.	PaaS (databases, web apps), IaaS
3	IBM SmartCloud	SaaS including data warehousing and analytics, business analytics engine, business process management, financial modelling, payment systems, medical analysis, social media analysis, transportation management, medical analytics, SQL databases, NoSQL databases, mobile integration.	SaaS (social media analysis), PaaS (databases, web apps), IaaS
4	Microsoft Azure	Windows or Linux virtual machines, messaging, scheduling, SQL databases, NoSQL databases, mobile integration.	PaaS (databases), IaaS
5	Rackspace Cloud	High bandwidth networking, virtual machines, data storage, process load balancing.	IaaS

Analysis of Technical Requirements and Alternatives versus Commercial Cloud Providers

Representative current cloud solution offerings from commercial vendors include but are not limited to the following: Amazon Web Services (AWS), IBM SmartCloud, Microsoft Azure, Google Cloud Platform, and Rackspace Cloud Servers. Considering the data management, experimentation requirements and the strategic issues, the question arises, “Are the IaaS and

⁸ In computing, FLOPS (for FLoating-point Operations per Second) is a measure of computer performance, useful in fields of scientific calculations that make heavy use of floating-point calculations. For such cases, it is a more accurate measure than the generic instructions per second. Computers capable of performing greater than 1 Giga FLOPS are termed as supercomputers.

⁹ VMware, Inc. is a software company that provides cloud and virtualization software and service.

PaaS from these selected vendors sufficient for hosting and maintaining the FUSEnet SaaS and PaaS?” A summary of the cloud providers and the offering are described in Table 3.

Analysis of SaaS Technical Alternatives

FUSEnet consists of four unique technical emergence software applications. Current cloud providers are not in the business of providing this niche capability. Cloud providers offer more general SaaS services such as Enterprise Resource Planning (ERP), general accounting, medical, and financial applications for managing business administration operations. If FUSEnet were to be employed on a 3rd party cloud, unique, domain-specific expertise would be required to operate and manage the FUSEnet software applications.

Analysis of PaaS Technical Alternatives

FUSEnet consists of several framework and middleware solutions combined with math-based libraries that are unique to network and text analysis. With the exception of IBM SmartCloud, current cloud providers are not in the business of exclusively providing this niche capability. Cloud providers offer more general PaaS software such as databases, email, and web servers. The features of the network and social analytics tools in SmartCloud should be further evaluated.

Analysis of IaaS Technical Alternatives

FUSEnet is operated in a secured, cloud environment at the Data Computing Center at ORNL. It currently operates on the hardware infrastructure described above. This FUSEnet hardware was performance tested to determine its disk I/O (input/output) throughput under various load conditions. Software programs were used to perform these tests at a low level or ‘raw’ I/O set of read and write tests and at the application layer with tests that simulated application disk usage. From these initial test results and further repeated testing, the FUSEnet disk I/O was optimized for handling the volume and type of data used in the system. Further tests were performed to compare FUSEnet with another commercial cloud offering, which demonstrated similar or better performance for FUSEnet depending on the operating conditions selected. Currently, the FUSEnet storage system is in its second generation as a result of these performance tests and evaluations. The FUSEnet software and data can be operated on 3rd party (IaaS) environments that can meet the overall system requirements as follows:

- Handle big data that is mixed structured and unstructured and continuously growing.
- Protect selected data and apps (commercial, proprietary) that remain in the cloud.
- Rapidly deploy software solutions to the data.
- Provide virtualization for operating systems including common Linux distributions, Windows and Mac OS.
- Rapidly ingest data into the system.
- Provide the computing performance involving big data analytics software services.
- Provide an easy-to-use big data analytics platform.
- Provide high-performance big data storage and retrieval up to 500 TBs and continue to scale.
- Provide robust, state-of-the-practice cyber security.

In general, commercial firms are advised to consider strategic issues with regards to cloud scope, service levels, and deployment needs. For the FUSEnet environment, Table 4 summarizes these strategic concerns.

The overall need for a secured FUSEnet environment involves the capability to employ software services, such as the analytics described earlier, that uses the data within the FUSEnet cloud, but cannot copy the data out of the cloud. FUSEnet is equipped with custom

middleware software within the PaaS called a Data Diode that monitors activities and prevents the exfiltration of data. Thus, the commercial and proprietary data is protected from being taken outside the FUSEnet enclave (Abercrombie, MacIntyre, & Schlicher, 2011). The Data Diode involves a change to the Linux distro (distribution)¹⁰ so that an IaaS provider must approve the customer to host their own virtualized and configurable operating system (MacIntyre, Paul, & Schlicher, 2011).

Table 4. Strategic issues for the FUSEnet environment.

	<i>Strategic Issue</i>	<i>Description</i>	<i>Assessment for FUSEnet</i>
1	Cloud Scope – what is the design to meet the need?	Identifies the availability, performance, and security needs; sufficient and planned computing power, storage, and bandwidth.	FUSEnet is monitored daily and reported monthly with the current operational stats: Availability: 99.8%; CPU usage: 12-18%; Memory usage: 56-65%; Storage usage: 69%. FUSEnet is installed with a Data Diode that protects against data exfiltration of its repository. FUSEnet is a virtual environment with separated computing enclaves. Each user or user group within an enclave has the freedom to compose and perform their needed computational research without directly impacting other users.
2	Service Levels	Identifies the expected workload, admin support, service delivery needs, timing and I/O response.	FUSEnet Test and Evaluation (T&E) simulates heavy end-user loading. This is measured to be an increase of 5-10% of the daily load. For its initial usage, FUSEnet could simultaneously host 3-4 heavy end-users loading. The Admin support is at two levels: operating system and the virtual layer through VMware.
3	Deployment Needs	Identifies the integration needs with infrastructure services.	FUSEnet operates on VMware that isolates the PaaS from dependencies on the hardware and the Operating System. The current FUSEnet system, including the number of cores, performance of the cores, memory, and the Isilon storage, is a proven baseline for simultaneously hosting 3-4 heavy end-user loading.

Discussion and Conclusions

This paper addresses science policy with a method and a technique to assess research, increasing its value to the US national scientific community by making available a computing environment to support repeatable scientific big data experimentation of world-wide scientific literature. The computational capability ensures the integrity, availability and confidentiality of new technologies and new technical knowledge. This will position scientific investigators (academic, commercial, and government) with an advantage to address the technical and political challenges all three entities face. FUSEnet offers this unique capability and this paper describes a computing environment necessary to support repeatable experimentation, and recommends a system that is housed at the ORNL Data Center in order to provide value to investigators from a variety of sources while adhering to recently mandated Data Management Planning.

¹⁰ A Linux distribution (often called a distro for short) is an operating system made as a collection of software based around the Linux kernel and often around a package management system

Acknowledgments

We thank colleagues and other reviewers for their assistance and helpful comments. This research was supported by the Intelligence Advanced Research Projects Activity (IARPA) via the Department of Energy (DOE). This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the U.S. Department of Energy. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DOE, ORNL, or the U.S. Government.

References

- 2 CFR 215 - Uniform Administration Requirements for Grants and Agreements With Institutions of Higher Education, Hospitals, and Other Non-Profit Organizations (OMB Circular A-110) (2012).
- Abercrombie, R. K., MacIntyre, L. P., & Schlicher, B. G. (2011). *Protection of Data in Virtual and Physical Computing Environments* (Invention Disclosure Number: 201102659, DOE S-Number: S-124,217). Oak Ridge: Oak Ridge National Laboratory.
- Azoulay, P. (2012). Research efficiency: Turn the scientific method on ourselves. *Nature*, 484(7392), 31-32.
- Babko-Malaya, O., Hunter, D., Amis, G., Meyers, A., Thomas, P., Pustejovsky, J., et al. (2013, May 8-10). *Characterizing Communities of Practice in Emerging Science and Technology Fields*. Paper presented at the 2013 International Conference on Social Intelligence and Technology (SOCIETY).
- Bettencourt, L. A., Kaiser, D., Kaur, J., Castillo-Chávez, C., & Wojick, D. (2008). Population modeling of the emergence and development of scientific fields. *Scientometrics*, 75(3), 495-518.
- Brock, D. C., Babko-Malaya, O., Pustejovsky, J., Thomas, P., Stromsten, S., & Barlos, F. (2012, November 2-4). *Applied Actant-Network Theory: Toward the Automated Detection of Technoscientific Emergence from Full-Text Publications and Patents*. Paper presented at the Association for the Advancement of Artificial Intelligence (AAAI) Fall Symposium Social Networks and Social Contagion, Arlington, VA.
- Collins, D. (2014). *Knowledge Article: Data Management*. Oak Ridge: Oak Ridge National Laboratory.
- Data management plan. (2014). Retrieved June 21, 2015 from: http://en.wikipedia.org/wiki/Data_management_plan
- Dhar, V. (2013). Data Science and Prediction. *Communications of the ACM*, 56(12), 64-73.
- Dissemination and Sharing of Research Results. (2010). Retrieved June 21, 2015 from: <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>
- Fürstenau, H., & Rambow, O. (2012). *Unsupervised induction of a syntax-semantics lexicon using iterative refinement*. Paper presented at the First Joint Conference on Lexical and Computational Semantics.
- Huang, M.-H., Huang, W.-T., Chang, C.-C., Chen, D.-Z., & Lin, C.-P. (2014). The Greater Scattering Phenomenon Beyond Bradford's Law in Patent Citation. *Journal of the Association for Information Science and Technology*, 65(9), 1917-1928.
- Intangible property - Code of Federal Regulations 2 CFR 200.315(2014).
- Kuhn, T. S. (1970). *The Structure of Scientific Revolutions* (3rd ed.). Chicago: University of Chicago Press.
- MacIntyre, L. P., Paul, N. R., & Schlicher, B. G. (2011). *Data Diode* (Copyright Document Number 90000008). Oak Ridge: Oak Ridge National Laboratory.
- Michaelis, J. R., McGuinness, D. L., Chang, C., Luciano, J. S., & Hendler, J. (2012). *Applying Multidimensional Navigation and Explanation in Semantic Dataset Summarization*. Paper presented at the 11th International SemanticWeb Conference (ISWC 2012).
- Murdick, D. A. (2011). Foresight and Understanding from Scientific Exposition (FUSE). Retrieved June 21, 2015 from: <http://www.iarpa.gov/index.php/research-programs/fuse>
- Otto, E., & Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*, 28(6), 441-453.
- Readron, S. (2014). Text-mining offers clues to success. *Nature*, 509, 410.
- Statement on Digital Data Management. (2014). *DOE Office of Science Funding Opportunities* Retrieved June 21, 2015 from: <http://science.energy.gov/funding-opportunities/digital-data-management/>
- Tian, W., & Zhao, Y. (2015). Chapter 1 - An Introduction to Cloud Computing. In W. Tian & Y. Zhao (Eds.), *Optimized Cloud Resource Management and Scheduling* (pp. 1-15). Boston: Morgan Kaufmann.

Is Italy a Highly Efficient Country in Science?

Aparna Basu¹

¹*aparnabasu.dr@gmail.com*

Formerly at CSIR-NISTADS CSIR National Institute of Science Technology and Development Studies,
New Delhi (India)

Abstract

In an earlier study on measuring national efficiencies in the production of scientific papers and patents of several developed and developing countries (Basu, 2013; 2014a), we found that Italy has the highest efficiency in the production of papers. While this has not gone unnoticed in the literature (Daraio and Moed, 2011) they have taken it as an ‘overcompensation effect’ and an indication of decline. By examining the work of several authors, we find instances where the information put forward, when taken together, support our findings – that Italy has a high efficiency in scientific publication but only an average efficiency in patenting. We note that Italy’s profile along a host of parameters is quite distinct with respect to the OECD average (DeJaeger, 2012). Using a typology of countries based on their publication and patenting efficiencies (Basu, 2014b) we infer that Italy is not one of the countries that have shifted national priorities from publications to patents, like USA, Japan, Germany, or Korea.

Conference Topic

Science policy and research assessment

Introduction

According to Hollanders and Soete, investment on R&D (GERD) is a correlate of development (Hollanders and Soete, 2010). Developed countries have higher GERD shares as compared to GDP shares, the Gross Expenditure on R&D (GERD) being the expenditure on the creation of new knowledge. Countries that have increased R&D expenditures, such that GERD share/GDP share tends to or exceeds unity, are on the path of development. How do increased investments of resources translate into outputs? Do developed countries make more efficient use of their resources? Efficiency of scientific productivity at the national level has been considered earlier by several authors (May, 1997; Rousseau, 1998; King, 2004, Vinkler, 2005, 2008; Shelton, 2008; Leydesdorff & Wagner, 2009; Wendt et al., 2012), who also point out difficulties in making cross-national comparisons. Primarily, they have dealt with publications and citations as compared to research expenditure or GNP and have considered mostly European countries, the US, Japan and China. Rousseau has considered both publications and patents. More recently, Shelton and Leydesdorff have also considered outputs such as patents and number of graduates in addition to papers, using regression models to predict outputs for a given set of inputs (Shelton and Leydesdorff, 2011). Some papers that have used different techniques such as Data Envelopment Analysis (DEA) to study national research productivity and efficiency are Rousseau (1998), Sharma and Thomas (2008) and Lee (2005). According to Hu et al., who used the distance function approach, intellectual property rights protection, technological cooperation among business sectors, knowledge transfer between business sectors and higher education institutions, agglomeration of R&D facilities, and involvement of the government sector in R&D activities significantly improve national R&D efficiency (Hu, et al., 2014)

In our earlier study on the efficiencies of nations in the production of scientific outputs with respect to inputs such as manpower and expenditure in science, we found significant variation in their efficiencies (Basu, 2013, Basu, 2014a). In particular, we noted that the efficiency of production of papers with respect to both expenditure on R&D (GERD) and manpower were the highest for Italy. This fact has not gone unnoticed the literature on Italian science. Daraio

and Moed (2011) did an extensive study on manpower, research expenditure, publications and citations to compare Italy with other productive EU countries. They called Italy “a Cathedral in the desert”, but at the same time chose to focus on other factors to argue that Italian science was in decline. Our attempt here is to see if there were other indications in the literature which could have pointed to the fact of Italy’s high efficiency, but were missed at the time.

Data and Methodology

Data on scientific papers and patents is taken from the SCI-Expanded and USPTO for the years 2008 and 2007. (The data and analysis are from our earlier papers (Basu, 2013, 2014a) and reproduced here for convenience.) Restricting to the USPTO, the United States Patent Office, gives a bias in favour of the USA termed as the ‘home advantage’. Ideally data from some of the other major patent databases such as the European Patent Office EPO should be included in the analysis. However for this preliminary study we have only considered the USPTO.

The Gross Domestic Product GDP and Gross Expenditure on Research and Development GERD for the years 2002 and 2007, are both adjusted to Purchasing Power Parity (PPP) in order to make local investments comparable across countries. Manpower is measured in terms of Full Time Equivalents (FTEs) engaged in R&D. Data is obtained from the UNESCO Science Report 2010 (UNESCO, 2010).

The share of GERD and the share of GDP are shown for a selected set of developing and developed countries Table 1. The GERD/GDP share is an indicator of development (Hollanders & Soete, 2010).

Table 1. GERD and GDP shares of selected countries (2002 and 2007).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	GDP	GDP	GERD	GERD	GERD	GERD	GERD
	share	share	share	share	share/	share/	share/
Country	2002	2007	2002	2007	GDP share	GDP share	GDP share
					2002	2007	2007-2002
EU	25.3	22.5	26.1	23.1	1.03	1.03	0.00
USA	22.5	20.7	35.1	32.6	1.56	1.57	0.01
China	7.9	10.7	5	8.9	0.63	0.83	0.20
Japan	7.4	6.5	13.7	12.9	1.85	1.98	0.13
Germany	4.9	4.3	7.2	6.3	1.47	1.47	0.00
India	3.8	4.7	1.6	2.2	0.42	0.47	0.05
France	3.7	3.1	3.9	3.4	1.05	1.10	0.04
UK	3.7	3.2	3.9	3.4	1.05	1.06	0.01
Italy	3.3	2.8	2.2	1.9	0.67	0.68	0.01
Brazil	2.9	2.8	1.6	1.8	0.55	0.64	0.09
Russia	2.8	3.2	2.0	2.0	0.71	0.63	-0.09
Mexico	2.1	2.3	0.5	0.5	0.24	0.22	-0.02
Korea	2.0	1.9	2.8	3.6	1.40	1.89	0.49
Canada	2.0	1.9	2.4	2.1	1.20	1.11	-0.09
Australia	1.3	1.2	1.3	1.4	1.00	1.17	0.17

Table 2 shows the manpower and GERD figures (in FTE’s and billion \$ PPP) together with the output of papers in the Science Citation Index-Expanded using fractional counts, and patents in the USPTO.

Table 2. Manpower, GERD, Papers and Patents for selected countries.

Country	GERD \$bnPPP	Manpower (FTE's)	Papers SCI-E	Patents USPTO
Australia	15.36	87,140	28,313	1,516
Brazil	20.20	133,266	26,482	124
Canada	23.96	139,011	43,539	3,806
China	102.40	1,423,380	104,968	7,362
France	42.89	215,755	57,133	3,631
Germany	72.24	290,853	76,368	9,713
India	24.79	154,827	36,261	741
Italy	22.12	96,303	45,273	1,836
Japan	147.90	709,974	74,618	33,572
Korea	41.30	221,928	32,781	6,424
Mexico	55.90	37,930	8,262	81
Russia	23.40	451,213	27,083	286
Spain	19.34	130,896	35,739	363
UK	41.04	261,406	71,302	4,007
USA	398.00	1,425,550	272,879	81,811

Definitions

To define efficiency we have considered some inputs and outputs in the science system, and their ratio output/input. The inputs have been taken as the expenditure and manpower in research. The outputs are scientific patents and papers published by the nations. For two inputs and two outputs there are four possible components of efficiency (Basu, 2013).

The efficiency for paper production for each country has two values $EE(Pap)$ and $ME(Pap)$, defined for expenditure and manpower as,

$$\text{Expenditure Efficiency } EE(Pap) = \text{Papers}/\text{GERD} \quad (1)$$

$$\text{Manpower Efficiency } ME(Pap) = \text{Papers}/\text{Manpower} \quad (2)$$

where GERD is the national expenditure on R&D (in PPP), and the manpower is in terms of full time equivalents in R&D (FTE's).

The efficiency for patent production also has two values $EE(Pat)$ and $ME(Pat)$,

$$\text{Patent Expenditure Efficiency } EE(Pat) = \text{Patents}/\text{GERD} \quad (3)$$

$$\text{Patent Manpower Efficiency } ME(Pat) = \text{Patents}/\text{Manpower} \quad (4)$$

While papers and patents are homogeneous entities, GERD is made up of several components such as HERD, BERD, GOVERD, which are the expenditures on the Higher Education sector, the business sector and the government sector. Each of these components contributes in a different way to output of papers and patents. For example, expenditure in the business sector is expected to give rise to patents rather than papers, Higher education and government expenditures give rise to primarily papers, while defence expenditure, which is part of expenditure in the government sector does not produce many papers or patents. While this indicates that questions of efficiency are more complex than what has been considered here, in the present study we will use GERD as a single homogeneous entity.

Analysis

In Table 1 we see the inputs made by a set of selected countries in the years 2002 and 2007 to R&D (GERD), expressed as a share. A country is taken to be a developed country if its share of GERD is higher than its share of GDP (GERD share/GDP share >1; Hollanders & Soete, 2010). Using this criterion we see from Table 1 that in both 2002 and 2007 the EU as a whole, USA, Japan, Germany, France, UK, Australia and Korea had GERD share/GDP share >1, and would be termed developed countries. We note that Italy is missing from this list, although it is a part of the EU. It is listed along with China, India, Brazil, Mexico, Russia for which GERD/GDP is less than 1. The data indicates that expenditure on R&D in Italy is lower than would be expected for a developed country.

A plot of Expenditure efficiency and Manpower efficiency in the production of scientific papers shows that Italy has the highest efficiency in both directions (Fig. 1). This implies for the amount of money invested and manpower deployed in the R&D system, Italy has the highest efficiency. This observation makes Italy and its science system an interesting object of study.

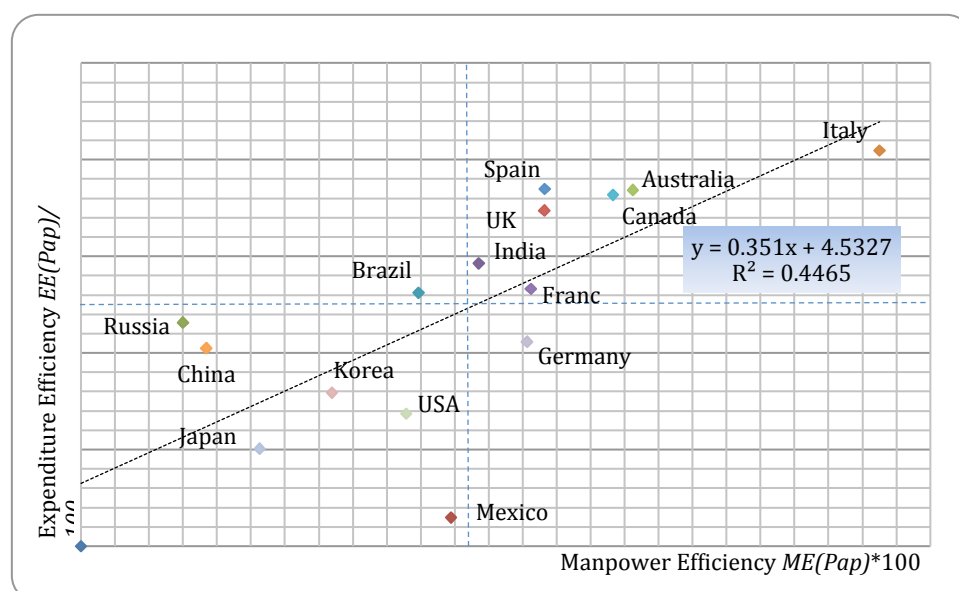


Figure 1. Efficiency of paper production with respect to expenditure EE(Pap) and manpower ME(Pap). Note that Italy scores very high on both dimensions.

For patent production we have the corresponding quantities EE(Pat) and ME(Pat) calculated using Eqns 3 and 4, and plotted in Figure 2. Here we note that USA, Japan are at the highest level in patenting efficiency, while Germany, Korea and Canada are at a medium level. UK and Australia are just above average and Italy and France are somewhat above the average (blue dotted lines) on manpower efficiency ME(Pat) but below average on expenditure efficiency EE(Pat). China, India, Spain, Mexico, Brazil and Russia are below average in patenting efficiency.

The high degree of collinearity ($R^2=0.9$) in the graph suggests that manpower and expenditure are correlated, which is not surprising since a large fraction of the expenditure usually goes toward salaries. This is also true to some extent of the efficiencies of paper production (Fig. 1).

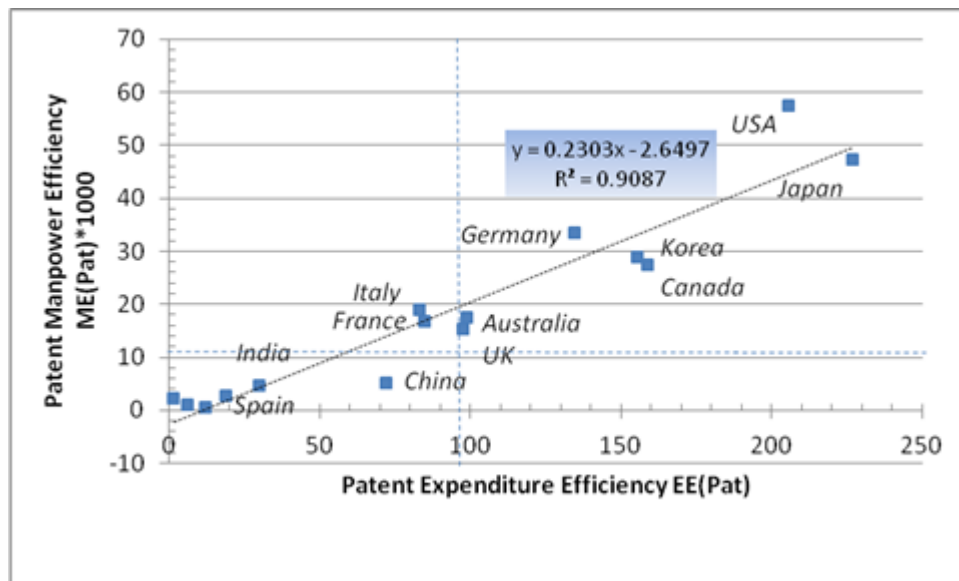


Figure 2. Efficiency of patent production with respect to expenditure EE(Pat) and manpower ME(Pat). Countries in left corner are Russia, Mexico and Brazil.

It should be emphasized that since there are 4 dimensions, two-dimensional graphs give only a partial picture of the similarity of profiles of the different countries.

The case of Italy

The case of Italy is somewhat unusual because of the very high values of efficiency of paper production with respect to both manpower and expenditure (Fig. 1). While this has not been explicitly stated in the literature, it is possible that there were indications of it in the work of others (Daraio & Moed, 2011; Foland & Shelton, 2010). Our attempt will be to trace such instances that support our finding. Firstly, we consider expenditure and recall that Italy had GERD share/GDP share less than unity, which categorises it with developing countries (Table 1). In Figure 3 we look at the GERD values of some countries (OECD data, 2012). Among a set of European countries together with US and Japan, Italy has the lowest value of the input GERD as a percentage of GDP. Since efficiency is the ratio of output to input, a low value of input raises efficiency. Spain also has a low value of expenditure, which makes its publication efficiency with respect to expenditure high. However its publication efficiency with respect to manpower is low (Fig. 1)

In terms of the business component of GERD (BERD) and the Government expenditure (GOVERD) the same trend prevails (Figs. 4 & 5) showing that Italy has almost the lowest values among these countries. This has also been noted in Daraio and Moed (2011).

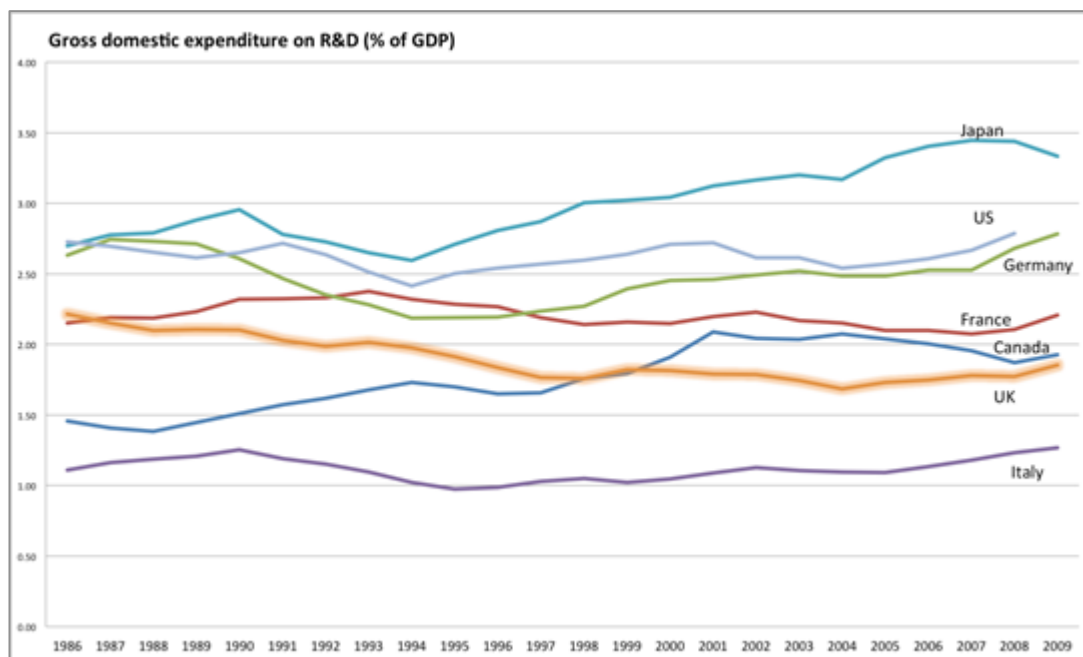


Figure 3. Gross domestic Expenditure on R&D (Source: OECD data, 2012).

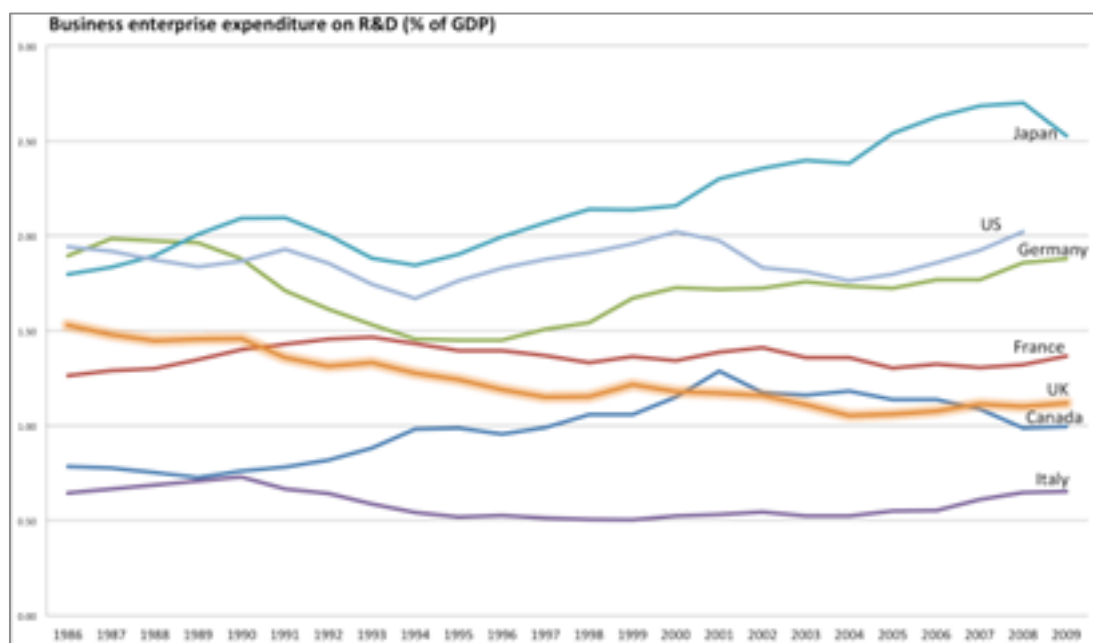


Figure 4. Business Enterprise data BERD (Source: OECD data, 2012).

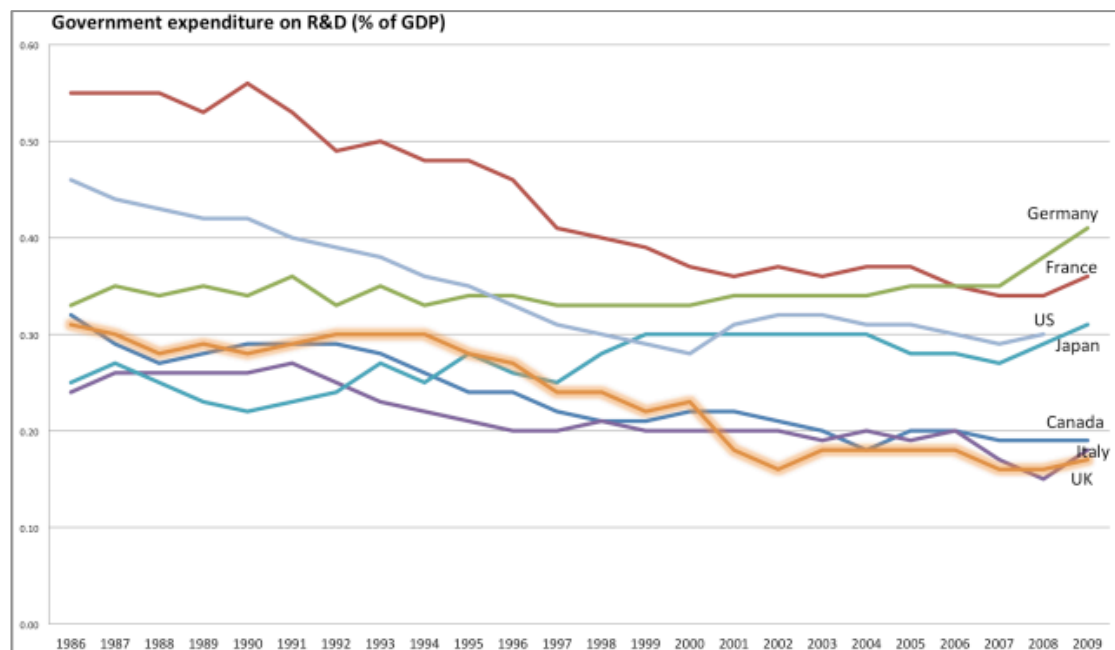


Figure 5. Government Expenditure on R&D, GOVERD (Source: OECD data, 2012).

Figures. 3-5 show that Italy has one of the lowest values of R&D expenditure as a share of GDP among all the countries shown. It also had the lowest expenditure on military R&D spending, a sector not expected to produce many papers or patents, as seen from Figure 6 reproduced from Foland and Shelton (2010).

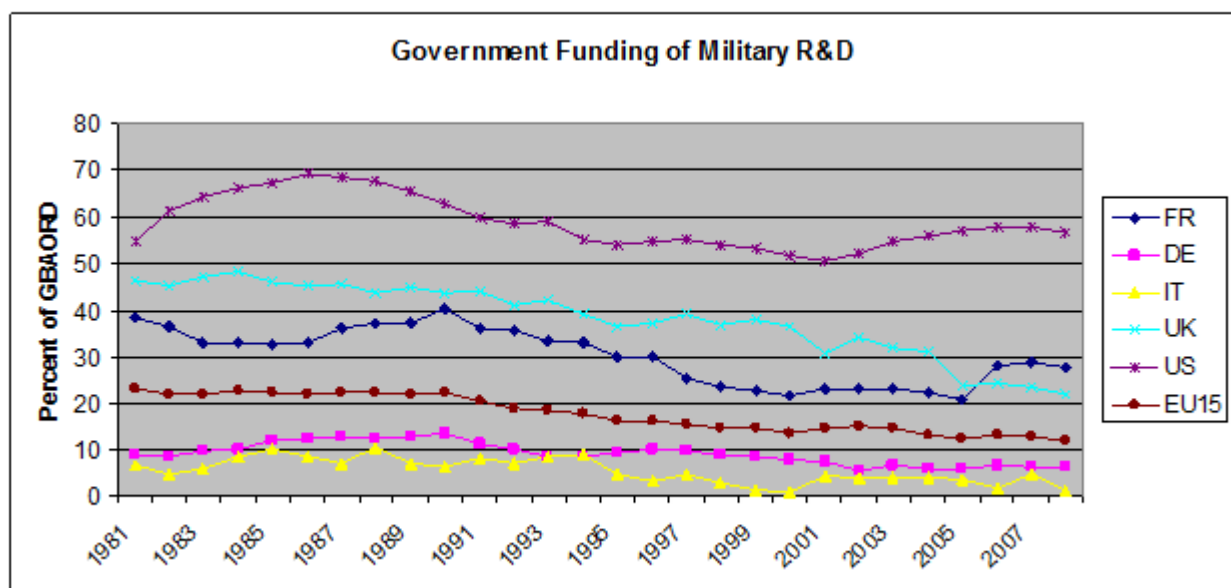


Figure 6. Government Funding of military R&D, showing that Italy has one of the lowest military spending (Source: Foland and Shelton, 2010).

At the same time, in a graph by the same authors showing growth rates of published papers for different countries, it is clearly seen that Italy had the highest growth rate over two successive decades (Fig. 7). Thus it would appear that there has been an efficiency increase with respect to expenditure for Italy, both due to lowered expenditure on R&D as well as increases in publication output.

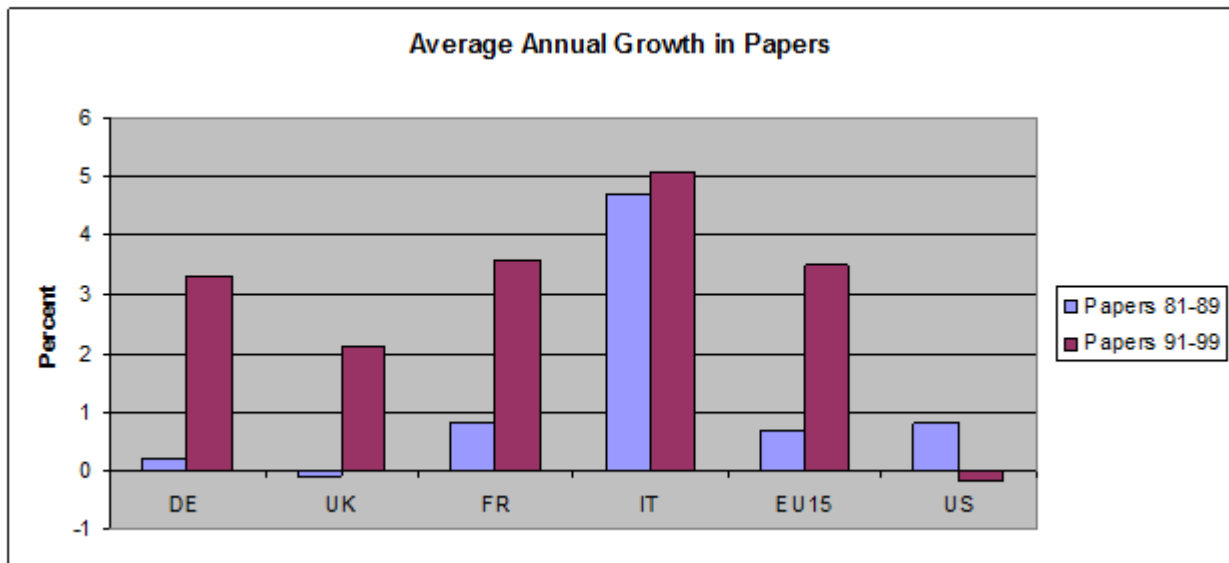


Figure 7. Average annual growth in papers, for a few selected countries (Source: Foland & Shelton, 2010).

Finally, we find more detailed information in a series of country profiles created by DeJaeger (2010) for 39 OECD countries and some developing countries. From Italy's profile a comparison of Italy's outputs with other countries shows some interesting points. DeJaeger's profile of Italy is reproduced below (Fig. 8).

The GERD is low, about 1.26% of GDP, about half the OECD average and more in line with the R&D intensity of emerging economies, as seen earlier.

The manpower values are also lower than OECD average. At the same time the output of papers is on par with the average output of the group of OECD countries. This would give Italy a higher efficiency of publication with respect to manpower as compared to the average. Daraio and Moed (2010) also note in their paper that Italy's publications grew in the period 1980-2009, till it had the highest publication output per researcher amongst other European countries (*see Figure 8 in Daraio and Moed; they however, they prefer to use papers per thousand population as an index instead, and predict a decline for Italy based on a lack of correlation between citation impact and manpower values.*)

In brief, while the number of researchers per thousand total employment is low compared to the average, Italy's output of papers per million population is on par with the average of the other countries, making its efficiency high for publications (Fig. 8). Triadic patents per million population is very low compared to other countries (Fig. 8), which coupled with low values of expenditure and research manpower lead to a medium value for patenting efficiency (Fig. 2).

Another point of interest is the high percentage of foreign funding in GERD as compared to other countries. DeJaeger (2012) notes that internationalization in Italy is high. About 41% of scientific articles and 13% of PCT patents were produced with international collaboration. In 2009, industry funded 44% of GERD, Government funding was 42% and 9% was funded from abroad. Regarding international collaboration Daraio and Moed find that Italy's share of internationally co-authored bilateral papers is lower than other OECD countries and their role (vis a vis first authorship) is like the developing countries (Fig. 4 in Daraio & Moed, 2011). From Figure 8 we also see that Italy has a higher number of foreign co-inventors as compared to other countries. It is possible that foreign funds apply to these sectors.

In summary Italy appears to be a country, which has achieved a high efficiency of publication of papers funded with low funds a substantial part of which is from foreign sources. Its

expenditure in the business sector is also low, but its patenting is close to average again indicating medium efficiency.

One limitation of our study is that citations have not been considered in the definition of efficiency. Even though Italy's citations appear to be favorable in some studies (Aspen Report, 2012, Dario & Moed, 2011), it is possible that considering citations would give a different picture. Other caveats common to most bibliometric studies refer to the use of publications as homogeneous units without reference to disciplinary biases in productivity and efficiency, difficulties in comparing expenditures (should one use Purchasing Power Parity, PPP \$?), as well as manpower due to differing conventions in different countries (Wendt, et al, 2011).

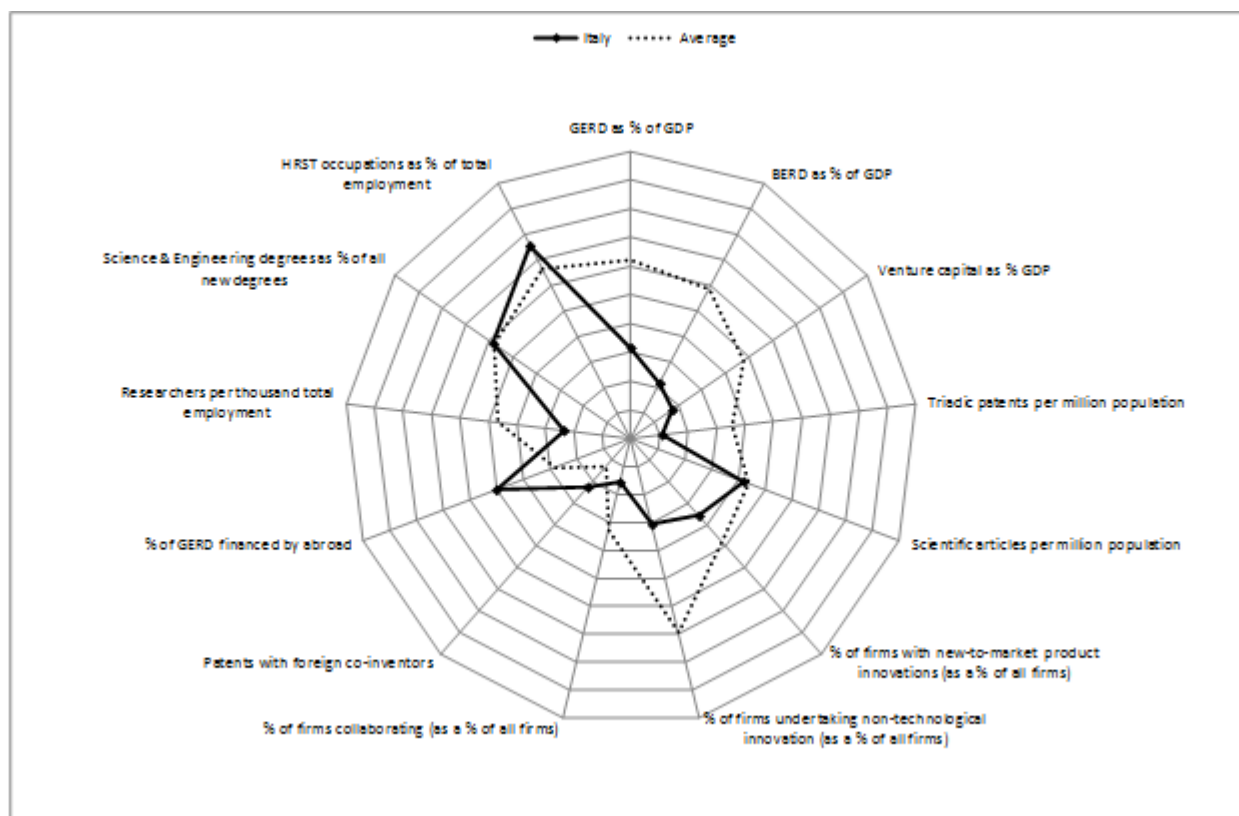


Figure 8. Italy's position vis-à-vis OECD countries on several parameters related to science (Source: De Jaeger, 2012).

Discussion

Efficiency of different countries in the production of papers and patents with respect to manpower and expenditure were calculated by us to obtain a national comparison of R&D efficiency. Unlike many earlier studies on efficiency that included only OECD and other developed countries and Japan and China, we have included several developed and developing countries (see also Basu, 2014, a, b). It was found that Italy had the highest efficiency in the production of papers as compared to the developed and developing countries (Fig. 1). We concluded that Italy has an unusual profile which though noticed in the literature, has not been further investigated (Aspen Report, 2012; Daraio & Moed, 2011). In Italy, research expenditure as a fraction of GDP was found to be low, not only in comparison with other OECD countries, but actually in line with developing countries as noted by us here. At the same time scientific articles per million population are on par with the average OECD value (Fig 8). Italy's expenditure on the military R&D sector is also low (Foland & Shelton; 2010). This may be contrasted with the US where 50% of the government expenditure goes to

the military. Since the defence sector is one that does not produce papers or patents, this gives an advantage to Italy in the computation of efficiency in terms of scientific publication output. In other words, Italy spends much less of its GDP on R&D as compared to other developed countries, at the same time achieving the same rate of publications per million population as other OECD countries (OECD figures; Figure 8). Dario and Moed (2011) refer to this as the ‘Cathedral in the desert’.

Research manpower as a proportion of total employees is also much lower than the average OECD value, but science degrees are at the average OECD value (Fig. 8). We note here that the OECD makes a distinction between researchers and human resources in S&T (HRST) where the latter would include technical staff. HRST figures as a proportion of total employees in Italy are much higher than average OECD values (Fig. 8). The possible implication of this is that in Italy, the mix of research staff (academic, research, technical) may be different compared to other countries, with a higher component of technical staff. Since a large part of research expenditure goes towards salaries, and technical staff is likely to be less well paid, this may be a contributing factor toward economy in research expenditure. This conjecture needs to be validated by further research.

All of these features where output is average but inputs are low contribute to high efficiency, which is what we have observed in the case of Italy. In case it should appear that high efficiency in the case of Italy is only because of low inputs, it should be pointed out that growth in the output of papers was the highest for Italy over two successive decades (Foland & Shelton, 2010; Daraio & Moed, 2011). Another possible factor in achieving higher levels of publication than expected from low investments in R&D could be international funding and high collaboration. A substantial part of GERD in Italy comes from foreign sources (Fig. 6).

However, the number of patents are low, not only in the USTPO as seen in our study but also for Triadic patents as seen in the country profiles by De Jaeger. Since the expenditure outlay is also low in the business sector which contributes more to patents (BERD; Figure 4), the efficiency in patenting given by their ratio is close to average (Fig. 2). At the same time the number of foreign co-inventors is high, almost double the OECD value (Fig. 6).

In addition to the observations above regarding possible explanations for the high efficiency in science and relatively lower efficiency in patenting in Italy, we refer to our recent paper on a typology of countries based on research efficiency (Basu, 2014b). According to Basu, as national priorities shift from publications to patents as they appear to have done, fuelled by large increases in the business component of GERD, countries have witnessed a fall in publications (not only through the “displacement effect” due to the rise of China) coupled by a rise in patent efficiency. Countries that have moved in this direction are the USA, Japan, and Germany. Italy apparently has not made this transition, and is characterized by very low levels of investments by the business sector and low efficiency in patenting, but a high efficiency in publication. (Shelton and Leydesdorff have used expenditure in the government and business sectors and shown their relation to different outputs, Shelton & Leydesdorff 2011).

While Shelton and Ali (2011) have noted other countries like Turkey, Greece, Poland and Slovakia as being scientifically efficient, Italy appears to have been missed. Daraio and Moed (2011) in their detailed study ‘Is Italian science declining?’, observed that Italy had the highest productivity per researcher, and among the lowest levels of R&D expenditure for a selected set of EU countries, (for the period around 2007-2008), but instead of regarding it as efficiency, they argued on the basis of lower levels of foreign collaboration and publication output per 1000 inhabitants and detailed policy analysis that Italy was on the verge of a decline in science. They attributed the performance to an ‘overcompensation effect’, and state that the “the productivity of the system is often used in the political debate to justify a further cut in spending”, underlining their apprehensions.

In summary, it appears that Italy has produced over 3% of the world's papers and shown the highest growth rate in two decades (amongst EU countries) with a modest outlay (in line with less developed countries), both in terms of expenditure and manpower in a demonstration of high efficiency in basic science. Of greater concern is the fact that Italy is only average in patenting efficiency, and falls below OECD averages in BERD, venture capital, technological firms undertaking innovative activities or with technological products to market. On the international front, it has much higher contribution to GERD from foreign funds and has almost twice as many co-inventors as compared to other OECD countries.

Acknowledgements

The author acknowledges a grant under the Emeritus Scientist Scheme of the Council of Scientific and Industrial Research, New Delhi (2010-2014), and thanks anonymous referees for their comments.

References

- Albuquerque, E. (2005). Science and Technology systems in less developed countries. In H.Moed, W. Glanzel, U.Schmoch (Eds.) *Handbook of Quantitative Science and Technology Research* (pp. 759-778) Kluwer Academic Publishers.
- Aspen Institute Italia. (2012). Research in Italy, Strengths and Weaknesses.
- Basu, A. (2013). Efficiencies in national scientific productivity in terms of manpower and funding in science, in *Proceedings of the 14th International Society for Scientometrics and Informetrics (ISSI) Conference*, Vienna, July 15-19, 2013: (pp. 1954-1956) http://www.issi2013.org/Images/ISSI_Proceedings_Volume_II.pdf
- Basu A. (2014a). The Albuquerque model and efficiency indicators in national scientific productivity with respect to manpower and funding in science, *Scientometrics*, 100(2), 531-539.
- Basu, A. (2014b). A typology of countries based on efficiency of publication and patenting with respect to manpower and expenditure, in Noyons, E. (Ed.) Context Counts – Pathways to Master Big and Little Data, Proceedings of the Science and Technology Indicators Conference, Leiden
- Daraio, C. & Moed, H.F. (2011). Is Italian science declining? *Research Policy*, 40(10), 1380-1392.
- DeJaeger, Nils. (2010). OECD Science, Technology and Industry Outlook 2010. <http://www.oecd.org/sti/inno/oecdsciencetechnologyandindustryoutlook2010.htm>
- DeJaeger, Nils. (2012). OECD Science, Technology and Industry Outlook, p. 328 <http://www.oecd.org/sti/oecdsciencetechnologyandindustryoutlook2010.htm>
- Eurostat. (2012). Statistics Explained, European Commission. http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/R_%26_D_expenditure
- Foland, P. & R.D. Shelton (2010). Why is Europe so efficient at producing scientific papers, and does this explain the European Paradox? *11th International Conference on S&T Indicators, Leiden, Sept. 10, 2010*.
- Hollanders, H. & L. Soete (2010). The growing role of knowledge in the global economy. In *UNESCO Science Report 2010*, UNESCO Publishing.
- Hu, J.-L., Yang, C.-H., & Chen, C.-P. (2014). R&D Efficiency and the national innovation system: an international comparison using the distance function approach. *Bulletin of Economic Research*, 66, 55-71. doi:10.1111/j.1467-8586.2011.00417.x
- King, D.A. (2004). The scientific impact of nations, *Nature*, 430, 311-316.
- Lee, H. Y. & Park, Y. T. (2005). 'An international comparison of R&D efficiency: DEA approach', *Asian Journal of Technology Innovation*, 13, 207-22.
- Leydesdorff, L. & Wagner, C. (2009). Macro-level indicators of the relations between research funding and research output, *Journal of Informetrics*, 3(4), 353-362.
- May, R.M. (1997) The scientific wealth of nations, *Science*, 7 February 1997: 793-796.
- OECD (2011). "Business R&D", in OECD Science, Technology and Industry Scoreboard 2011, OECD Publishing. http://dx.doi.org/10.1787/sti_scoreboard-2011-18-en
- Rousseau, S. & Rousseau, R. (1998). The scientific wealth of European nations: Taking effectiveness into account, *Scientometrics*, 42(1), 75-87.
- Sharma, S. & Thomas, V.J. (2008). Inter-Country R&D Efficiency Analysis: An Application of Data Envelopment Analysis, *Scientometrics*, 76, 483-501.
- Shelton, R.D. (2008). Relations between national research investment and publication output: Application to an American Paradox. *Scientometrics*, 74(2), 191-205.

- Shelton R.D. & Leydesdorff, L. (2011). Publish or patent: Bibliometric evidence for empirical trade-offs in national funding strategies, *Journal of the American Society for Information in Science and Technology*, 63(3), 498-511.
- UNESCO Science Report (2010). UNESCO Publishing.
- Vinkler, P. (2005). Science indicators, economic development and the wealth of nations. *Scientometrics*, 63, 417-419.
- Vinkler, P. (2008). Correlation between the structure of scientific research, scientometric indicators and GDP in EU and non-EU countries. *Scientometrics*, 74(2), 237-254.
- Wendt K., Aksnes, D.W, Sivertsen, G., et al. (2012). Challenges in Cross-National Comparisons of R & D Expenditure and Publication Output, In *Proceedings of 17th International Conference on Science and Technology Indicators*, 2(0167) 826-834.

Performance Assessment of Public-Funded R&D Organizations Working on Similar Research Streams: A Multinational Study

Debnirmalya Gangopadhyay¹, Santanu Roy² and Jay Mitra³

¹ *debn4u@gmail.com*

National Institute of Science Technology and Development Studies (NISTADS), K.S. Krishnan Marg, New Delhi- 110012 (India)

² *sroy@imtdubai.ac.ae*

Institute of Management Technology (IMT), Dubai International Academic City, Dubai (United Arab Emirates)

³ *jmitra@essex.ac.uk*

Essex Business School, University of Essex Wivenhoe Park, Colchester, Essex CO4 3SQ (United Kingdom)

Abstract

The subject of deriving a measure of efficiency of public-funded organizations (primarily not-for-profit organizations) and of ranking these efficiency measures have been major subjects of debate and discussion. In the present study, the methodology of data envelopment analysis (DEA) has been used to analyze the relative performances of public funded R&D organizations across multiple countries working in similar research streams with multiple measures of inputs and outputs. The keywords highlighting the major research areas in the field of non-metrology conducted by National Physical Laboratory (NPL) in India were utilized to select the global comparators working in similar research streams. These global comparators were three R&D organizations located in the USA and one each located in Germany and Japan. The relative efficiencies of the organizations were assessed with variables such as external cash flow (ECF) earned, technologies transferred, publications and patents as outputs and grants received from the parent body and scientific personnel as inputs. The study indicates suggested measures and a set of targets to achieve the best possible performance for NPL and other R&D organizations.

Conference Topic

Science Policy and Research Assessment

Introduction

Public funded research and development (R&D) organizations utilize public money either through government-supported research programs or other public supported activities. These organizations carry out scientific research, deliver technological services to the society and play a fundamental role in an increasingly knowledge-based society ushering in innovations necessary for the development of a competitive industrial system. Research and innovation have become strategic resources and assets to foster competitive national economies (Coccia, 2005). The ability to attract, develop and retain high quality scientific and technical manpower as well as self-sustenance by means of minimizing its dependence on state funding assume vital importance as it impacts delivery that not only addresses national needs but also ensures traction on a global scale.

Globally, public R&D organizations are currently striving to improve their performance as a result of enhanced competition due to liberalization and globalization, increasing demands on the existing resources and being accountable for optimum allocation of these resources. As the R&D process utilizes scarce resources, it becomes crucial to assess the efficiency of this process (Sharma & Thomas, 2008). In the recent past government efficiency concerns have increased, more so in the light of diminishing funds (Gupta et al., 2000). The emerging demand for evaluating the performance of R&D organizations is the result of relentless growth in global competition (Tassey, 2009). However, the provision of quality information

to decision makers through a performance measurement system assumes criticality in such a scenario (Cook et al., 1995).

One major problem in evaluating the efficiency of public institutions is the lack of a good estimate of the production function. The breakthrough came in the research work undertaken by Charnes, Cooper and Rhodes (1978), the first paper using the technique of data envelopment analysis (DEA), even though they never named it that way. The present study makes an attempt to assess the relative efficiency of the National Physical Laboratory (NPL), a constituent establishment of the Council of Scientific and Industrial Research (CSIR), India, with five selected global comparators working in the same research streams located in three countries - the USA, Japan and Germany. Finally, suggesting measures have been proposed highlighting a set of targets to achieve the best possible performance for those R&D organizations, which are less efficient.

Literature Review

It is difficult to measure the performance of an R&D organization because the nature of these organizations and the functions these organizations perform are complex, risky, and uncertain. As opined by Chiesa and Masella (1996), Bremser and Barsky (2004), Loch and Tapper (2001), Brown and Svenson (1998), and Jain and Triandis (1997), it is difficult to identify, measure and compare the performance of R&D organizations. Further, researchers have found it difficult to identify the various outputs/inputs as multiple parameters are involved in the system. As per the existing literature, there exists only a few studies that have been conducted on performance measurement of R&D organizations (Roy, Mitra & Debnath, 2013; Garg et al., 2005).

R&D Output

Considering individual firms as the sample of their study, Pandit, Wasley and Zach (2011) consider R&D as an input to the innovation process and measures the productivity of a firm's innovative activities in terms of the number and the quality of patents. They argue that both of these variables are measures of innovation output or success, and proxy for the economic value of innovation. Chen, Hu and Yang (2011) suggest a multi-dimensional measurement schema including patents, royalties and licensing fees and journal articles. In their study on R&D and the national innovation system, Hu, Yang and Chen (2014) compare R&D efficiency among 24 nations during 1998-2005. In their multiple input-output framework, the input variables are R&D expenditure stock and R&D manpower and the output variables are patents, scientific journal articles, and royalty and licensing fees. Considering public research institutes, Matsumoto et al. (2010) have carried out case studies on market-impact creation outputs from the National Institute of Advanced Industrial Science and Technology, and have modelled R&D output generating economic impact along four stages – R&D output, technology transfer, commercialization, and market impact. This is in line with Roy et al.'s (2003) earlier study where a model to measure the effectiveness of research units was conceptualized. Likewise, research carried out by Laliene and Sakalas (2014) and Agostino et al. (2012) refer to the development of conceptual frameworks for R&D productivity assessment in public research organizations. Lee et al. (2011) have presented an R&D performance monitoring, evaluation and management system for national R&D to mirror not only short-term but also long-term R&D outcomes.

Methodology

Data envelopment analysis (DEA) as developed by Charnes et al. (1978) and extended by Banker et al. (BCC) (1984) has opened up new possibilities in evaluating the performances of many different kinds of entities (referred to as decision making units, DMU), engaged in

different activities and contexts (Cooper et al., 2004). DEA has been used widely to evaluate the performances of countries and regions (Rousseau and Rousseau, 1997, 1998), banks (Brockett et al., 1997), US air force wings (Charnes et al., 1985a), universities (Reichmann, 2004), Japanese manufacturing firms (Goto & Suzuki, 1989), journals (Lozano & Salmeron, 2005), R&D funding on education (Garg et al., 2005), etc. Publications and patents are used extensively to measure R&D efficiency and innovation (Pavitt, 1985). Evaluation of R&D efficiency could be advantageous to identify the better performers for benchmarking and choose better ways to improve efficiency highlighting areas of weakness (Sharma & Thomas, 2008). Charnes et al. (1985) have characterized a unit as influential if it is frequently used in the calculation of efficiency scores.

Researchers who have adopted the DEA methodology to evaluate performances of public research institutes include Rama Mohan (2005) and Roy, Mitra and Debnath (2013). Kim and Oh (2002) conducted a study on designing an R&D measurement system for Korean researchers. Wang et al. (2005) have developed extensive evaluation criteria for multidisciplinary R&D projects in China for ranking and rewarding. Roy et al. (2007) have earlier carried out a study on CSIR exploring the impact of age, research area, and rank on its scientific productivity, again using DEA as one of the methodologies.

Contextual Background of the Study

National Physical Laboratory (NPL), a premier institute of the Council of Scientific and Industrial Research (CSIR), India, has had a commendable track record of contributions and accomplishments since its inception and its scientists have received recognition for their contributions. Though maintenance and up-gradation of national standards of measurements remains the statutory responsibility of the organization, it is also involved in advanced non-metrology related research activities including engineering and electronic materials, material characterization, radio and atmospheric sciences, superconductivity and cryogenics.

A participatory workshop was conducted to diagnose NPL's R&D operations and to focus on aspects related to R&D performance. A particular research area (non-metrology) was selected for the purpose of the current analysis, and accordingly, the keywords, highlighting the organization's major research areas in this field, were utilized to shortlist global comparators. The keywords were searched in the SCOPUS database for a five-year period and global R&D organizations working on similar research streams were shortlisted. Five public R&D organizations were selected based on higher number of publications. These global comparators were the following:

- 1) National Institute for Materials Science, Japan (NIMS-JP, DMU-A),
- 2) National Renewable Energy Laboratory, USA (NREL-US, DMU-B),
- 3) Fritz Haber Institute of the Max Planck Society, Germany (FHI-DE, DMU-C),
- 4) National Centre for Atmospheric Research, USA (NCAR-US, DMU-D), and
- 5) Oak Ridge National Laboratory, USA (ORNL-US, DMU-E).

Data structure

The data regarding the inputs and outputs were collected for each DMU including NPL for a five-year period and are presented in Table 1. To ensure confidentiality, the exact period of the data cannot be revealed. Input variables considered in this study were: (1) grants received from the parent body, and (2) the number of scientific personnel (SP) whereas the output variables were: (1) business generated from the industry *i.e.*, external cash flow (ECF) earned, (2) technologies transferred (TT), (3) publications, and (4) number of patents filed.

The methodology to compare performance of any set of research institutes as suggested by Rama Mohan (2005) has been adopted in the present study. To illustrate the results on

efficiency assessment of public R&D organizations including NPL, one input variable and two output variables were considered at the same time.

Table 1. Input and output of different public R&D organizations (five year data).

<i>Public R&D Organization</i>	<i>Input</i>		<i>Output</i>			
	<i>Grants (Million USD)</i>	<i>Scientific Personnel (No.)</i>	<i>Technologies Transferred (no.)</i>	<i>Publication (No.)</i>	<i>Patents (No.)</i>	<i>ECF (Million USD)</i>
NIMS-JP - A	94	675	95	7480	195	20
NREL-US - B	141	307	53	2012	99	15
FHI-DE - C	72	206	1	1225	6	3
NCAR-US - D	185	310	5	2345	14	17
ORNL-US - E	107	1075	83	9144	90	23
NPL, India	47	216	3	1024	13	4

The DEAOS software was used for analysis. It analyzes relative performance of business units performing similar functions with an easy to use interface. It provides numerical and graphical output for easy interpretation and communication of results. Some of the key features of DEAOS are:

- The possibility to deal with 25 to ‘unlimited’ decision making units.
- Flexible facilities – importing from Excel file and direct entry of the data.
- Provides flexible input data management - possibility of addition and deletion of DMUs as well as rows and columns.
- Model input/output orientation selection.
- Provides a tabular scores report (with a variety of sorting methods) and a graphical summary.

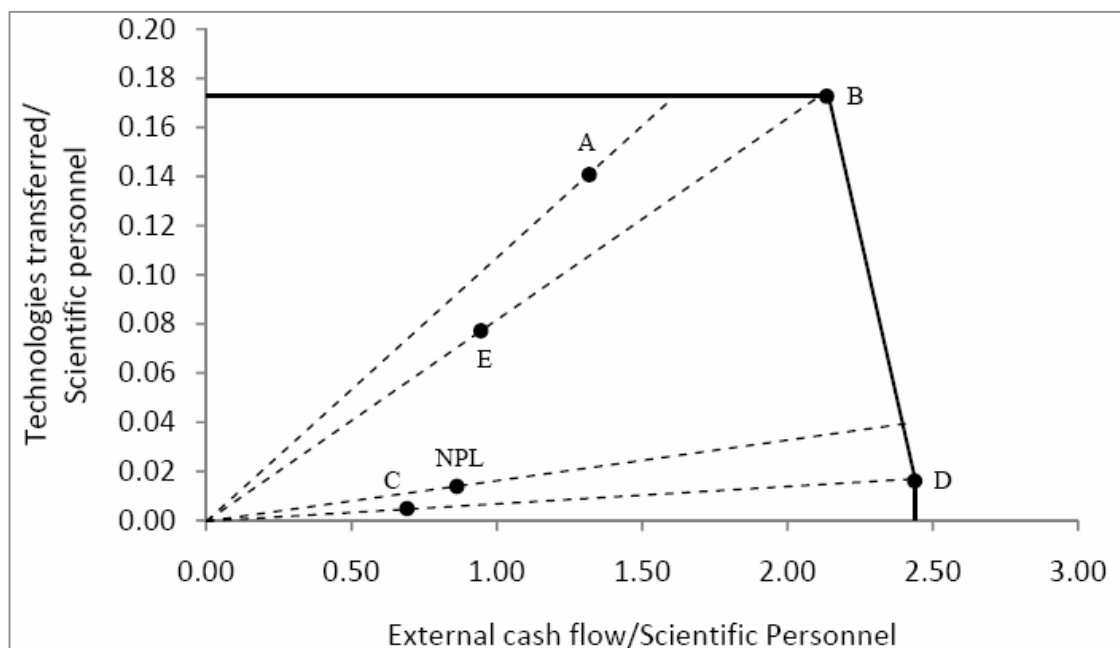
Results

ECF generated and technologies transferred vs. scientific personnel

Ratios were calculated for each organization (Table 2) along two dimensions viz., ECF generated per scientific personnel and technologies transferred per scientific personnel. Figure 1 clearly shows that NREL-US (DMU-B) and NCAR-US (DMU-D) are the best performers exhibiting 100% relative efficiency. The efficient frontier, which envelops NIMS-JP (DMU-A), FHI-DE (DMU-C), ORNL-US (DMU-E) and NPL, represents relative efficiency of those organizations. It is observed that NIMS-JP, FHI-DE, ORNL-US and NPL exhibited relative efficiencies of 82, 28, 45 and 36 % respectively. To enhance efficiency from 36 to 46%, NPL is assumed to increase the input-output ratios from the current level of 0.86 to 1.10 (ECF/scientific personnel) and 0.014 to 0.018 (technologies transferred/scientific personnel). An improvement target of 10 %, keeping input (scientific personnel) constant, can be achieved during the next year, if NPL is in a position to increase its ECF to 1.6 M USD and transfer at least 1 technology (Table 3).

Table 2. External cash flow (ECF) and technologies transferred vs. scientific personnel.

<i>Public Organization</i>	<i>R&D ECF / Scientific Personnel</i>	<i>Technology Transferred / Scientific Personnel</i>
NIMS-JP - A	1.32	0.14
NREL-US - B	2.13	0.17
FHI-DE - C	0.69	0.00
NCAR-US - D	2.44	0.02
ORNL-US - E	0.94	0.08
NPL, India	0.86	0.01

**Figure 1. ECF generated and technologies transferred vs. scientific personnel.**

Publications and patents vs. scientific personnel

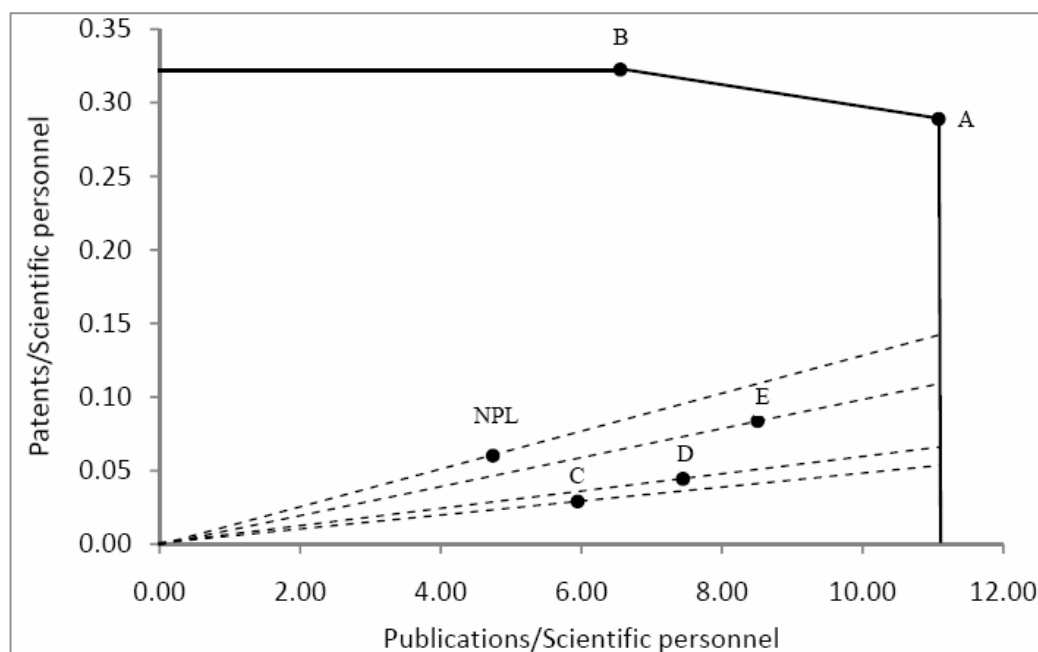
To assess the relative performance of the R&D organizations, publications per scientific personnel and patents per scientific personnel were calculated (Table 4) and graphically represented in Figure 2. NIMS-JP (DMU-A) and NREL-US (DMU-B) show best performance exhibiting 100% efficiency in generating sufficient number of publications and patents per scientific personnel. Performance was found higher in case of ORNL-US (DMU-E) (77%) and NCAR-US (DMU-D) (67%) whereas FHI-DE (DMU-C) (54%) and NPL (43%) perform moderately. However, NIMS-JP is the reference laboratory all the organizations. To achieve improved targets by 10% during the next year, NPL and FHI-DE each would require to publish 240 and 230 papers and 9 and 12 patents respectively (Table 5).

Table 3. Targets for the R&D organizations to improve efficiency by 10%

(Scientific personnel count remaining constant)		
<i>Public Organization</i>	<i>R&D ECF to earn (Million USD)</i>	<i>Technology to transfer</i>
NIMS-JP - A	6.8	12
FHI-DE - C	1.1	0.4
ORNL-US - E	5.1	19
NPL, India	1.6	0.8

Table 4. Publications and patents vs. scientific personnel

<i>Public Organization</i>	<i>R&D Publications / Scientific Personnel</i>	<i>Patents / Scientific Personnel</i>
NIMS-JP - A	11.08	0.29
NREL-US - B	6.55	0.32
FHI-DE - C	5.95	0.03
NCAR-US - D	7.44	0.04
ORNL-US - E	8.51	0.08
NPL, India	4.74	0.06

**Figure 2. Publications and patents vs. scientific personnel.****Table 5. Targets for the R&D organizations to improve efficiency by 10% (Scientific personnel count remaining constant).**

<i>Public Organization</i>	<i>R&D Publications</i>	<i>Patents</i>
FHI-DE - C	230	12
NCAR-US - D	347	23
ORNL-US - E	1204	96
NPL, India	240	9

ECF generated and technology transferred vs. grants

Next, relative efficiencies of the R&D organizations have been calculated along two outputs (ECF generated and technologies transferred) and one input (grants received from the parent body), (Table 6) and plotted in Figure 3. NIMS-JP (DMU-A) and ORNL-US (DMU-E) show best performance exhibiting 100% efficiency in generating sufficient amounts of ECF and number of technologies transferred per grants received. All the other organizations have ORNL-US in their reference set. To achieve efficiency by 10% during the next year, FHI-DE has to earn 1.5 M USD ECF and to transfer 7 technologies (Table 7).

Table 6. ECF earned and technologies transferred vs. grants received from parent body.

<i>Public R&D Organization</i>	<i>ECF / Grants</i>	<i>Technologies Transferred / Grants</i>
NIMS-JP - A	0.21	0.02
NREL-US - B	0.10	0.01
FHI-DE - C	0.04	0.00
NCAR-US - D	0.09	0.00
ORNL-US - E	0.21	0.02
NPL, India	0.09	0.00

Table 7. Targets for the R&D organization to improve efficiency by 10% (Grants received from the parent body remaining constant).

<i>Public R&D Organization</i>	<i>ECF to earn (Million USD)</i>	<i>Technology to transfer</i>
NREL-US - B	3	11
FHI-DE - C	1.5	7
NCAR-US - D	3.9	24
NPL, India	0.8	5

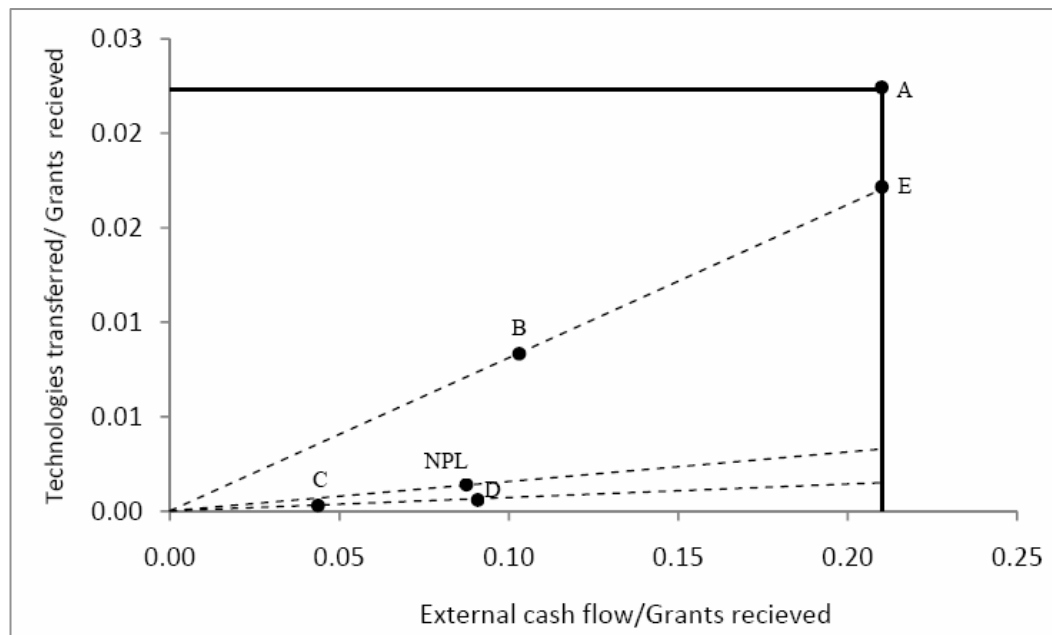


Figure 3. ECF generated and technologies transferred vs. grants received.

Publications and patents vs. grants

To assess the relative performance of the R&D organizations, ratios were calculated for publications per grants received and patents per grants received (Table 8) and graphically represented in Figure 4. NIMS-JP (DMU-A) and ORNL-US (DMU-E) show the best performance exhibiting 100% efficiency. NPL has both NIMS-JP and ORNL-US in its reference set whereas FHI-DE (DMU-C) and NCAR-US (DMU-D) relate only to ORNL-US whereas NREL-US (DMU-B) has only NIMS-JP in its reference set. To achieve efficiency by 10% during the next year, FHI-DE, NCAR-US and NPL have to increase their number of patents by a count of 7, 17 and 5 respectively from the current level (Table 9).

Table 8. Publications and patents vs. grants received from parent body.

<i>Public Organization</i>	<i>R&D Publication / Grants</i>	<i>Patent / Grants</i>
NIMS-JP - A	1.77	0.05
NREL-US - B	0.32	0.02
FHI-DE - C	0.38	0.00
NCAR-US - D	0.28	0.00
ORNL-US - E	1.89	0.02
NPL, India	0.48	0.01

Publication, patents, ECF generated and technology transferred vs. scientific personnel & grants

The relative efficiencies of R&D organizations on multi-input-multi-output six dimensional model keeping two inputs (*viz.*, scientific personnel & grants received) and four outputs (*viz.*, publication, patents, ECF generated and technology transferred) data have been calculated and the performance of each R&D organization under study is compared with that of every other one following the output oriented measure of efficiency at constant return to scale (CRS), variable return to scale (VRS) along with scale efficiencies (SE). The empirical analysis has been given in Table 10.

Table 9. Targets for the R&D organization to improve efficiency by 10 % (Grants received from the parent body remaining constant).

<i>Public Organization</i>	<i>R&D Publications</i>	<i>Patents</i>
NREL-US - B	1397	29
FHI-DE - C	617	7
NCAR-US - D	1575	17
NPL, India	399	5

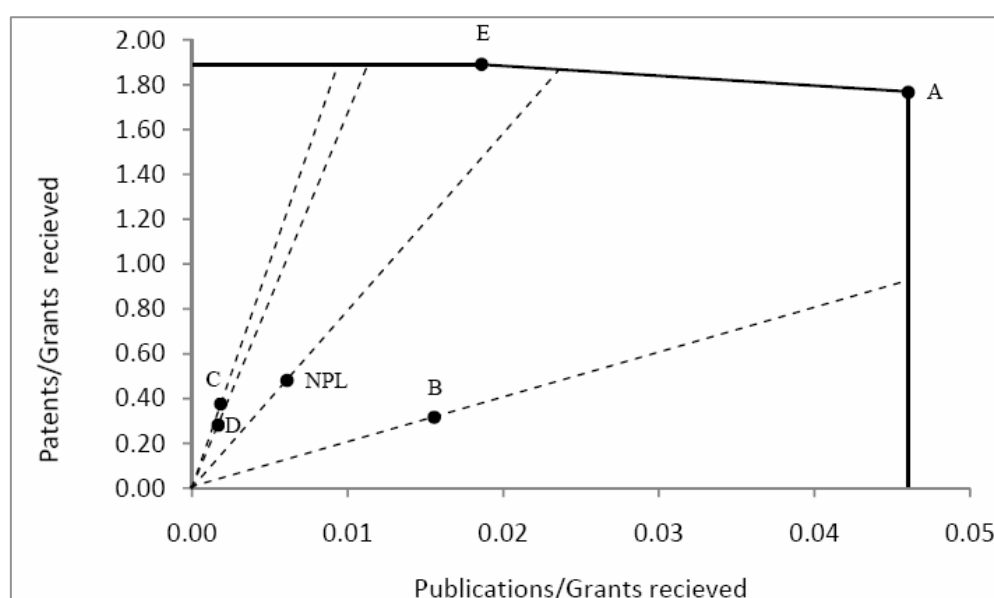
**Figure 4. Publications and patents vs. grants received.**

Table 10. Relative efficiency percentage of different public R&D organizations.

Public R&D Organization	ECF & TT/SP			Pub. & Pat/SP			ECF & TT/Grants			Pub. & Pat/Grants			Pub., Pat., ECF & TT/SP & Grants		
	CRS			CRS			CRS			CRS			CRS		
	VRS	SE	VRS	SE	VRS	SE	VRS	SE	VRS	SE	VRS	SE	VRS	SE	VRS
NIMS-JP - A	82	100	0.82	100	100	1.00	100	100	1.00	100	100	1.00	100	100	1.00
NREL-US - B	100	100	1.00	100	100	1.00	49	65	0.75	34	51	0.67	100	100	1.00
FHI-DE - C	28	100	0.28	54	100	0.54	21	25	0.84	20	27	0.74	54	100	0.54
NCAR-US - D	100	100	1.00	67	90	0.74	43	74	0.58	15	26	0.58	100	100	1.00
ORNL-US - E	45	100	0.45	77	100	0.77	100	100	1.00	100	100	1.00	100	100	1.00
NPL, India	36	94	0.38	43	85	0.51	42	100	0.42	26	100	0.26	57	100	0.57

Note: CRS: constant return to scale, VRS: variable return to scale SE: scale efficiency; (SE=CRS/VRS)

Technical efficiencies estimated under the CRS model are found to be less than the technical efficiencies coming from the more flexible VRS model. Under the CRS assumption, less average efficiency is found in case of FHI-DE (DMU-C) (54%) followed by NPL (57%) while under VRS, it was found that average technical efficiency score for all the DMUs is 100%, which implies that on an average DMUs could have used resources judiciously to produce the same amount of output. However, under the scale efficiency (SE), the average score is found to be 0.54 in case of FHI-DE and 0.57 in case of NPL, which indicate that on an average the actual scale of production has diverged from the most productive scale size. In SE, the score 1 indicates that the DMU is operating at the most efficient scale or optimal size whereas SE less than 1 would be due to decreasing returns to scale (over production) or increasing returns to scale (under production).

Discussion and Conclusions

Over the past three decades, a variety of approaches, parametric and non-parametric, have been developed to investigate the failure of producers to achieve the same level of efficiency (Kalirajan and Shand, 1999). DEA which offers a non-parametric alternative to parametric frontier production function analysis has two advantages over the econometric one in measuring productivity change (Grosskopf, 1986). First, it compares the states to the ‘best’ practice technology rather than ‘average’ practice technology as is done by econometric studies. Second, it does not require the specification of an ad hoc functional form or error structure. In DEA, the less-performing units need more inputs to produce the same amount of output (Andersen & Petersen, 1993). DEA produces a piecewise empirical extreme production surface which in economic terms represents the revealed best-practice production frontier (Charnes et al., 1994).

In this study, the performance of each R&D organization (here the DMU) under study is compared with that of every other one following the output oriented measure of efficiency at constant return to scale (CRS), variable return to scale (VRS) along with scale efficiencies (SE). DEA has been used to analyze the relative efficiencies of the public funded R&D organizations keeping one input and two outputs at a time and results have been demonstrated in four possible dimensions. Secondly, the relative efficiencies of R&D organizations on multi-input-multi-output six dimensional model keeping two inputs and four outputs data have also been calculated. Comparatively less efficiency of NPL (0.57) that is a cause for concern might be due to its lower efficiency in generating sufficient amounts of external cash flow, number of technologies assumed to be transferred to the industry per scientific personnel as well as number of papers published and patents filed per grants received from the parent body.

The significance of the work presented in the paper stems from the fact that this is perhaps the first multinational study of relative performance assessment of R&D organizations, all of whom work on similar research themes. Relative performance assessment of different R&D organizations have been ascertained in the past (Roy, Mitra & Debnath, 2013) but the R&D organizations in question were working on diverse research streams. The focus of the current study, therefore, seems much more relevant as absolute comparators were first identified and thereafter assessed in terms of their performance characteristics. The present work has opened up new avenues for further research in this area.

References

- Agostino, D., Arena, M., Azone, G., Molin, M.D., & Masella, C. (2012). Developing a performance measurement system for public research centres. *International Journal of Business Science and Applied Management*, 7(1), 43-60.
- Andersen, P. & Petersen, N. C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science*, 39, 1261-1264.
- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data development analysis. *Management Science*, 30, 1078-1092.
- Bremser, W.G. & Barsky, N. P. (2004). Utilizing the balanced scorecard for R&D performance measurement. *R&D Management*, 34(3), 229-237.
- Brockett, P. L., Charnes, A., Cooper, W. W., Huang, Z. M., & Sun, D. B. (1997). Data transformations in DEA cone ratio envelopment approaches for monitoring bank performances. *European Journal of Operational Research*, 98, 250-268.
- Brown, M.G. & Svenson, R.A. (1998). Measuring R&D productivity. *Research and Technology Management*, 41(6), 30-35.
- Charnes, A., Clark, T., Cooper, W. W., & Golany, B. (1985). A developmental study of data envelopment analysis in measuring the efficiency of maintenance units in the U.S air forces. *Annals of Operations Research*, 2, 95-112.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429-444.
- Charnes, A., Cooper, W. W., Lewin, A. Y., & Seiford, L. M. (1994). *Data Envelopment Analysis: Theory, Methodology and Applications*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Chen, C., Hu, J., & Yang, C. (2011). An international comparison of R&D efficiency of multiple innovative outputs: the role of the national innovation system. *Innovation Management, Policy and Practice*, 13(3), 341-360.
- Chiesa, V. & Masella, C. (1996). Searching for an effective measure of R&D performance. *Management Decision*, 34(7), 49-57.
- Coccia, M. (2005). A scientometric model for the assessment of scientific research performance within public institutes. *Scientometrics*, 65, 307-321.
- Cohen, W. M. & Levinthal, D. A. (1989). Innovation and learning: the two faces of R&D. *The Economic Journal*, 99, 569-596.
- Comanor, W. S. & Scherer, F. M. (1969). Patent statistics as a measure of technical change. *The Journal of Political Economy*, 77, 392-398.
- Cooper, W. W., Seiford, L. M., & Zhu, J. (2004). Data envelopment analysis: history, models and interpretations. In W. W. Cooper, L. M. Seiford & J. Zhu (Eds.) *Handbook on Data Envelopment Analysis*, (pp. 1-39). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Garg, K. C., Gupta, B. M., Jamal, T., Roy, S., & Kumar, S. (2005). Assessment of impact of AICTE funding on R&D and educational development. *Scientometrics*, 65, 151-160.
- Goto, A. & Suzuki, K. (1989). R&D capital, rate of return on R&D investment and spillover of R&D in Japanese manufacturing industries. *The Review of Economics and Statistics*, 71, 555-564.
- Grosskopf, S. (1986). The role of the reference technology in measuring productive efficiency. *The Economic Journal*, 96, 499-513.
- Gupta, A. K., Bhojwani, H. R., Kaushal, R., & Kaushal, H. (2000). Managing the process of market orientation by publicly funded labs: Case of CSIR. *R&D Management*, 30, 289-296.
- Hall, B. H., Griliches, Z., & Hausman, J. A. (1986). Patents and R&D: Is there a lag? *International Economic Review*, 27, 265-283.
- Hu, J. L., Yang, C.H., & Chen, C.P. (2014). R&D efficiency and the national innovation system: an international comparison using the distance function approach. *Bulletin of Economic Research*, 66(1), 55-71.

- Jain, R.K. & Triandis, H.C. (1997). *Management of Research and Development Organizations: Managing the Unmanageable*. 2nd Ed., New York: John Wiley & Sons.
- Kalirajan, K. P. & Shand, R. T. (1999). Frontier production functions and technical efficiency measures. *Journal of Economic Surveys*, 13, 149-172.
- Laliene, R. & Sakalas, A. (2014). Conceptual structure of R&D productivity assessment in public research organizations. *Economics & Management*, 19(1), 25-35.
- Lee, H., Kim, M.S., Yee, S.R., & Choe, K. (2011). R&D performance monitoring, evaluation and management system: a model and methods. *International Journal of Innovation and Technology Management*, 8(2), 295-313.
- Loch, C.H. & Tapper, U.A.S. (2001). Implementing a strategy-driven performance measurement system for an applied research group. *The Journal of Product Innovation Management*, 19, 185-98.
- Lozano, S. & Salmeron, J. L. (2005). Data envelopment analysis of OR/MS journals. *Scientometrics*, 64, 133-150.
- Matsumoto, M., Yokota, S., Naito, K. & Itoh, J. (2010). Development of a model to measure the economic impacts of R&D outputs of public research institutes. *R&D Management*, 40(1), 91-100.
- Pandit, S., Wasley, C.E., & Zach, T. (2011). The effect of research and development input and output on the relation between the uncertainty of future R&D performance and R&D expenditures. *Journal of Accounting, Auditing and Finance*, 26(1), 121-144.
- Pavitt, K. (1985). Patent statistics as indicators of innovative activities: possibilities and problems. *Scientometrics*, 7, 77-99.
- Rama Mohan, S. (2005). Benchmarking evaluation of performance of public research institutes using data envelopment analysis. *Journal of Scientific and Industrial Research*, 64, 403-410.
- Reichmann, G. (2004). Measuring university library efficiency using data envelopment analysis. *Libri*, 54, 136-146.
- Rousseau, S. & Rousseau, R. (1997). Data envelopment analysis as a tool for constructing scientometric indicators. *Scientometrics*, 40, 45-56.
- Rousseau, S. & Rousseau, R. (1998). The scientific wealth of European nations: taking effectiveness into account. *Scientometrics*, 42, 75-87.
- Roy, S., Mitra, J., & Debnath, R.M. (2013). Ranking R&D institutions of India: an application of DEA. *International Journal of Business Development and Research*, 1(2), 49-66.
- Roy, S., Nagpaul, P.S., & Mohapatra, P.K.J. (2003). Developing a model to measure the effectiveness of research units. *International Journal of Operations & Production Management*, 23(12), 1514-29.
- Sharma, S. & Thomas, V. J. (2008). Inter-country R&D efficiency analysis: an application of data envelopment analysis. *Scientometrics*, 76, 483-501.
- Tassey, G. (2009). Methods for assessing the economic impacts of government R&D. <http://www.nist.gov/director/prog-ofc/report03-1.pdf>.

Outlining the Scientific Activity Profile of Researchers in the Social Sciences and Humanities in Spain: The Case of CSIC

Adrián A. Díaz-Faes¹, María Bordons¹ Thed van Leeuwen² and M^a Purificación Galindo³

¹*adrian.arias@cchs.csic.es, maria.bordons@cchs.csic.es*

Quantitative Analysis in Science & Technology Group (ACUTE), IFS, Spanish National Research Council (CSIC), Albasanz 26-28, Madrid 28037 (Spain)

²*leeuwen@cwts.leidenuniv.nl*

CWTS-Centre for Science and Technology Studies, Leiden University, PO Box 905
2300 AX Leiden (the Netherlands)

³*pgalindo@usal.es*

Statistics Department, Salamanca University, Alfonso X El Sabio s/n, Salamanca 37007 (Spain)

Abstract

Scientific activity of Social Sciences and Humanities researcher's comprises an assorted set of publication channels such as books, book chapters and national and international journal articles. Since knowledge dissemination in the field is characterised by a greater use of national journals and local languages, international bibliographic databases do not offer a suitable coverage. This work pursues to draw a comprehensive picture of the publication behaviour of CSIC researchers in the Social Sciences and Humanities from a micro-level perspective. For this purpose, Web of Science and an internal CSIC database called 'ConCiencia' were used along with a set of indicators describing the activity profile of researchers as well as the prestige of publication channels. Differences in the publication pattern of researchers in SSH were explored, and the relationship between their research performance and personal features such as age, gender and professional rank were analysed. In the Humanities, researchers with higher academic rank and age showed greater activity in books and non-WoS articles, whereas in the Social Sciences, higher rank was related to internationally-oriented scientific publications and a more collaborative activity. Considering only WoS articles would shrink meaningfully the visibility of CSIC researchers.

Conference Topic

Science policy and research assessment

Introduction

Outlining the scholarly work of researchers in the Social Sciences and Humanities (SSH) is often regarded as a challenge in bibliometrics, since the predominant publication types in these fields are not well covered by large bibliographic databases such as Web of Science or Scopus (Hicks, 2004). At this point, it is quite clear that dealing with journal publications, it is not enough for the SSH (Archambault et al., 2006; Sivertsen & Larsen, 2012) remaining books and books chapters as a major communication channel, chiefly in the Humanities. Moreover, due to the more local orientation of research in the SSH, knowledge dissemination in the field is characterized by a greater use of national journals and local languages (van Leeuwen, 2013). On the other hand, even though there has been a certain trend to consider SSH as a whole, different behavior between both communities can be expected (Mañana-Rodríguez & Giménez-Toledo, 2013).

The aforesaid factors hinder the potential capacity of the traditional bibliometric analyses to provide a reliable picture of the scientific activity of the SSH researchers and the development of national or regional databases to obtain full coverage of publications in the SSH has been suggested (Martin et al. 2010). This type of database has been developed in some countries such as Norway, Denmark, Finland and Belgium (Flanders), motivated by the need to monitor the performance of university scholars and in line with the development of performance-based

funding of university research (Sivertsen, 2010). Studying the activity of SSH researchers in Spain is difficult, because there is not such a full coverage national bibliographic database, but it can be addressed at the institutional level because many institutions collect the scientific output of their researchers, mainly with evaluative purposes.

This study focuses on the scientific activity of SSH researchers at the Spanish National Research Council (CSIC), the largest public institution dedicated to research in Spain which makes up more than 4,000 researchers and 125 institutes spread all over the country. This work pursues to draw a comprehensive picture of the publication behaviour of CSIC researchers in SSH from a micro-level perspective. An assorted set of publication channels such as books, books chapters, international and national journal articles are considered and specific indicators to assess the prestige of the different publication channels are introduced. Differences in the publication pattern of researchers in SSH are explored, and the relationship between their research performance and personal features such as age, gender and professional rank are analyzed.

Methodology

This study analyses the scientific output of 268 active researchers in 2007 in the SSH area affiliated to the Spanish National Research Council (CSIC) and comprises both permanent researchers and postdoctoral research fellows. The time span under analysis is 2007-2011. Publications were collected from two different sources: Web of Science (WoS) (SSCI+AHCI+SCIE), which was used to download the more international articles; and an internal CSIC database called 'ConCiencia', to obtain other publication types not covered by WoS (books, books chapters and non-WoS journal articles). To cope with names inconsistencies and achieve a proper allocation of the publications to the researchers, different algorithms were used. A manual revision of the output collected, especially for the 'ConCiencia' database, was done. Based on the information retrieved, the following indicators were computed:

a) Activity profile of researchers

- % Books: proportion of books published by a researcher with regard to its total number of publications. In the same way, the next three indicators were calculated.
- % Book chapters.
- % WoS articles.
- % Non-WoS articles.
- Sum of publications: the total number of publications published by each researcher, including books, chapters in books and journal articles.
- Average number of authors/paper: this indicator measures the average number of authors per publication for the total output of a given researcher (WTI2, 2014).
- % International collaboration: share of the total output of each researcher co-authored with researchers affiliated with one or more foreign institutions.
- % English: proportion of a researcher's output published in English.

b) Prestige of publication channels

- Top books and chapters (pptop10% Books & Chapters): proportion of books and chapters of a given researcher published by the top 10% publishers according to the Scholarly Publisher Indicators Project (SPI) (Giménez-Toledo, Tejada-Artigas & Mañana-Rodríguez, 2013). This project describes the Indicator of Quality of Publishers according to Experts (ICEE), which is based on a quality assessment of publishers rated by Spanish researchers in a national survey.

- Proportion of papers in first quartile journals (Q1): share of papers published in the top 25% journals of the impact factor journal ranking by subject category (source: Journal Citation Reports).
- Proportion of papers in top non-WoS journals (pptop10% non-WoS articles): % of non-WoS papers published in top journals according to the Integrated Scientific Journal Classification (CIRC) (Torres-Salinas et al. 2010). CIRC is a proposal for a categorization of journals in SSH developed by a group of experts in bibliometrics in Spain. It distinguishes four categories of journals (A, B, C and D) according to their visibility measured integrating the results of different journal classifications and assessments tools. For the purposes of this paper, “top journals” are those included in the categories “A” and “B”.

Table 1. Impact indicators for the different types of publication channels.

<i>Type of publication channel</i>	<i>Indicators of impact/prestige</i>
WoS articles	Impact factor (25% top journals by impact factor)
Non-WoS articles	CIRC (categories A and B)
Books/Book chapters	SPI (10% top publishers by expert opinion)

c) Personal data: age, professional rank (P=postdoctoral research fellow, TS=tenured scientist, RS=research scientist and RP=research professor) and gender of researchers were provided by CSIC.

A preliminary inspection of the similarity between variables was explored by means of Multidimensional Scaling (MDS). Non-linear Principal Component Analysis (NLPCA) was used to explore the relationship between personal features of researchers and their performance. Statistical analyses were performed with SPSS (v.20).

Findings

A total of 268 researchers had at least one publication in the period 2007-2011. In the whole SSH area, men represented 59% of all researchers, average age of researchers was 50 years old, and half of the researchers were in the lowest scientific category (tenured scientist). Postdoctoral research fellows accounted for only 7% of researchers in the area. Small differences between the Humanities and Social Sciences can be observed in Table 2.

Table 2. Personal features and scientific rank of researchers in SSH.

		<i>Humanities</i> (N=192)		<i>Social Sciences</i> (N=76)		<i>Total</i> (N=268)	
Gender	Men	115	60%	42	55%	157	59%
	Women	77	40%	34	45%	111	41%
Rank 2007	Post-doc	12	6%	6	8%	18	7%
	Tenured scientists	98	51%	42	55%	140	52%
	Research scientists	46	24%	13	17%	59	22%
	Research professors	36	19%	15	20%	51	19%
Age		50 ± 9 (28-70)		49 ± 10 (32-70)		50 ± 9 (28-70)	

Note: age expressed as average ± standard deviation (min-max).

A total of 3,004 documents were published by CSIC researchers in SSH during 2007-2011. Differences between Humanities and Social Sciences in the main publication types used are observed: WoS articles predominate in the Social Sciences while book chapters are the most frequent publication channel in the Humanities (Table 3).

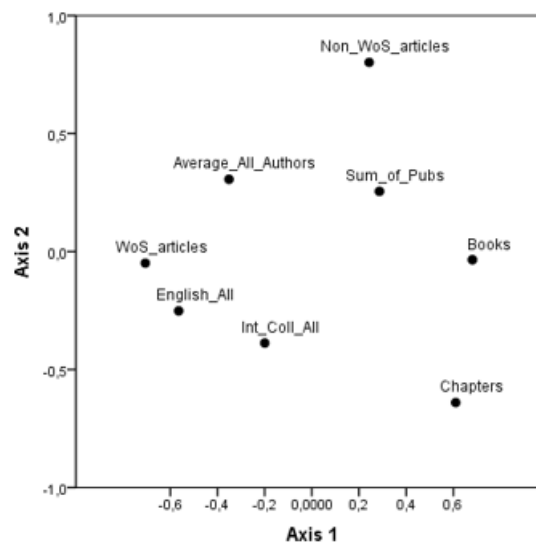
Table 3. Share of publication channels by area.

	<i>Books</i>	<i>Chapters</i>	<i>Non-WoS Articles</i>	<i>WoS Articles</i>	<i>Total</i>
Humanities	14% (397)	47 % (1,313)	26% (717)	13% (352)	2,779
Social Sciences	8% (65)	27% (214)	29% (227)	36% (289)	795
Total	13% (462)	43% (1,527)	26% (944)	18% (641)	3,574

Note: the total is higher than 3,004, because the publication count is made at the individual level.

Publication profile of researchers

A MDS was applied to the set of variables which make up the activity profile of researchers to reveal their underlying structure. In terms of similarity, the plot gives away greater levels of international collaboration and English-written publications for WoS articles. The patterns for the remaining publications types (books, chapters and non- WoS articles) seems to be mainly related to higher levels of productivity and being written in national languages (Figure 1).

**Figure 1. MDS for the scientific activity profile.**

The diversity of publication channels in the output of researchers is the norm in SSH. Around 1/3 of the researchers presented output of the four different types considered: articles covered by WoS, non-WoS articles, books and book chapters. Three and two types of publication channels were observed in 40% and 17% of the researchers respectively, while only 12% of researchers had results of a single type. Several differences between Social Sciences and Humanities can be put forward: researchers who disseminate research among the four different types of publication channels considered are more frequent in Humanities (36% vs 24%), while using only WoS-covered journals is more common among Social Sciences researchers (16% vs 4). Finally, it is interesting to remark that around 22% of Social Sciences researchers and 41% of those in the Humanities may remain invisible in Web of Science-based studies since they do not show any publication covered by this database.

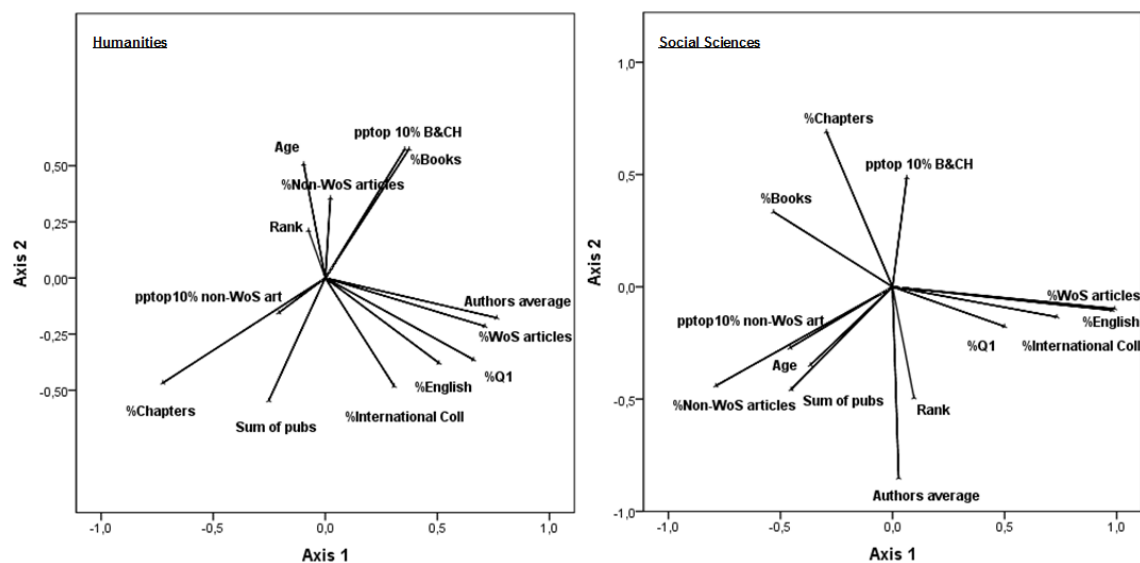
Research performance of scientists

Main statistics concerning research performance of scientists in SSH are shown in Table 4. A higher number of total publications is observed for researchers in the Humanities (15.1 vs 10.8), especially due to their high number of book chapters. Researchers in the Humanities exhibit a higher use of top publishers for books and chapters, while Social Sciences researchers present a greater share of articles in high impact factor journals.

Table 4. Description of the research performance of researchers in SSH.

	<i>Humanities</i>		<i>Social Sciences</i>	
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>
No. Books	2.1	2.5	0.9	1.0
No. Chapters	7.1	5.7	2.9	3.3
No. WoS Articles	1.9	4.3	3.9	4.1
No Non-WoS Articles	3.9	4.7	3.1	3.7
Sum of Publications	15.1	12.2	10.8	7.5
pptop10%_Books & Chapters	35.9	26.5	23.7	28.9
pptop10%_Non_WoS_Articles	32.8	35.3	37.3	37.9
% Q1 WoS Articles	12.9	29.7	33.4	36.5
Average number authors/publication	1.7	1.4	2.6	1.1
% International. collaboration	16.9	23.0	24.1	29.8
% English	14.0	19.0	38.6	32.7

To explore the possible relations between personal features of researchers and their performance NLPCA was used, which allows reducing a large number of variables to a smaller number of uncorrelated non-linear combinations of these variables with minimum loss of information (principal components). Two different studies are conducted, since researchers in Social Sciences and Humanities are analysed separately. Preliminary results concerning the plots of component loadings (two-dimensional solution) are shown in Figure 2.

**Figure 2. Component loadings in: a) Humanities; b) Social Sciences.**

Note: only researchers with 2 or more publications considered

Discussion and conclusions

At this point, some preliminary results can be pointed out in an attempt to provide a comprehensive picture of the activity of CSIC researchers in SSH from a micro-level perspective:

- Taking into account only WoS articles would shrink meaningfully the visibility of CSIC researchers in SSH, in particular in the Humanities.
- Different constraints of the ‘ConCiencia’ system are identified. More rigour in the input of data (carried out by researches themselves) as well as in the cleaning and validation processes (by the institution) would be advisable.

- In the Humanities, researchers who hold a higher rank and age present greater activity in books and non-WoS articles. However, a high number of total publications is apparently not associated to a higher rank.
- In the Social Sciences, a higher academic rank is associated to internationally-oriented scientific publications (high share of WoS articles) as well as a high productivity (high number of publications) and collaborative activity (high number of co-authors).
- Differences between the Social Sciences and Humanities are observed, but even within each of these fields different typologies of researchers according to their publication pattern, collaboration practices and international/national orientation may exist. These factors are being explored at present.
- Although our study focuses on four different types of academic output, it is still not comprehensive, since it does not consider the non-scholarly literature, which may have an important societal impact.

Acknowledgments

Adrián A. Díaz-Faes is granted with a JAE predoctoral fellowship by the Spanish National Research Council (CSIC).

References

- Archambault, E., Vignola-Gagne, E., Cote, G., Lariviere, V. & Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics*, 68(3), 329–342. doi:10.1007/s11192-006-0115-z.
- Giménez-Toledo, E., Tejada-Artigas, C. & Mañana-Rodríguez, J. (2013). Evaluation of scientific books' publishers in social sciences and humanities: results of a survey. *Research Evaluation*, 22(1), 64–77. doi:10.1093/reseval/rvs036
- Hicks, D. (2004). The four literatures of social science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems* (pp. 473–496). Dordrecht, The Netherlands: Kluwer Academic. doi:10.1007/1-4020-2755-9_22.
- van Leeuwen, T.N. (2013). Bibliometric research evaluations, Web of Science and the Social Sciences and Humanities: a problematic relationship? *Bibliometrie - Praxis und Forschung*, 2013, 1-18. Retrieved June 15, 2015 from: <http://www.bibliometrie-pf.de/article/viewFile/173/215>
- Mañana-Rodríguez, J. & Giménez-Toledo, E. (2013). Scholarly publishing in social sciences and humanities, associated probabilities of belonging and its spectrum: a quantitative approach for the Spanish case. *Scientometrics*, 94, 893-910. doi:10.1007/s11192-012-0838-y
- Martin, B., Tang, P., Morgan, M., Glänzel, W., Hornbostel, S., Lauer, G., et al. (2010). Towards a bibliometric database for the social sciences and humanities—A European scoping project. Research report produced for DFG, ESRC, AHRC, NWO, ANR and ESF.
- Sivertsen, G. (2010). A performance indicator based on complete data for the scientific publication output at research institutions. *ISSI Newsletter*, 6(1), 22–28.
- Sivertsen, G. & Larsen, B. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: an empirical analysis of the potential. *Scientometrics*, 91(2), 567–575. doi:10.1007/s11192-011-0615-3.
- Torres-Salinas, D., Bordons, M., Giménez, E., Delgado, E., Jiménez, E. & Sanz, E. (2010). Clasificación integrada de revistas científicas (CIRC): propuesta de categorización de las revistas en ciencias sociales y humanas. *El Profesional de la Información*, 19(6), 675-683.
- WTI2 (2014). Scholarly publication patterns in the social sciences and humanities and their relationship with research assessment. *Science, Technology & Innovation Indicators 2014. Thematic paper 2*, 1-26 Retrieved June 15, 2015 from: http://dialogic.nl/documents/other/sti2_themepaper2.pdf

A Bibliometric Assessment of ASEAN's Output, Influence and Collaboration in Plant Biotechnology

Jane G. Payumo¹ and Taurean C. Sutton²

¹ jane.payumo@kaust.edu.sa

4700 King Abdullah University of Science and Technology, Thuwal, 23955-6900 (Saudi Arabia)

² taurean.sutton@wsu.edu

1610 NE Eastgate Blvd., Suite 650, Washington State University, Pullman, WA, (United States)

Abstract

This research uses 10-year (2004-2013) publication and citation data related to plant biotechnology to assess the research performance, impact, and collaboration of member states of the ASEAN in plant biotechnology. Findings indicate increased scientific output of ASEAN countries in plant biotechnology as well as increased research collaborations by individual member states and with international partners throughout the 10-year period. The nature of collaboration by ASEAN is linked with the status of economic development of each country. Domestic and international collaborations are strong and are increasing through the years, regional collaboration on the other hand is found to be limited. This limited regional partnership can be a concern for the region's goal of economic integration. Further studies using bibliometric data analysis is suggested for policy diagnosis in plant biotechnology cooperation, knowledge flows, and effect of plant biotechnology research in economic development between ASEAN countries.

Conference Topic

Bibliometrics and research evaluation

Introduction

The Association of Southeast Asean Nations (ASEAN) has declared biotechnology as the main area of cooperation in science and technology. ASEAN, a regional association composed of 10 countries namely: Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar (Burma), Philippines, Singapore, Thailand, and Vietnam, considers plant biotechnology as the next pillar of regional economic growth (Hautea & Escaler, 2004; Erbisch & Maredia, 1998) and the answer to their food security needs. If ASEAN will continue to invest in plant biotechnology in the next years, it will be beneficial to have information on the current state of research and collaboration for strategic direction setting. This research drawing on bibliometric data, hence, will add to understanding the level and nature of collaboration, including research performance of ASEAN countries in plant biotechnology. This is relevant for ASEAN policy makers in charge of setting direction and designing strategies for research cooperation, and planning research investments, especially on biotechnology, at the country and regional levels.

Methodology

This research is based on 2004-2013 publications in plant biotechnology authored and co-authored by 10 member states of ASEAN. The data were extracted from Elsevier's Scopus database, the world's largest abstract and citation database of peer-reviewed literature (Elsevier B.V., 2014). Different keyword combinations were used to locate plant biotechnology-related publications guided by the glossary of biotech terms by the U.S. National Institute for Food and Agriculture (NIFA, 2014) and the National Agricultural Library Agricultural Thesaurus (National Agricultural Library, 2014). Additional filter was then set according to affiliation country to include only the publications published by the 10

ASEAN countries. No filter was set for the type of publication; all document type, namely: article, review, conference paper, short survey, note, editorial, letter, book chapter, book, and article in press were included. This research also highlights the use of a home-grown open-source ‘publication parser’ tool (Sutton, 2013); this tool was useful in parsing extracted files from Scopus for analysis of various indicators of interest at the country, institutional, and individual levels. The methodology, including interpretation of the different indicators, builds on best practices on indicators research that have been developed throughout the years (Moed, Glänzel, & Schmoch, 2004).

Results and Discussion

Publication output and citation impact

During the 10-year period (2004-2013), ASEAN researchers produced an overall total of 7,907 papers related to plant biotechnology; this output has increased 15% per year. These publications were written by more than 13,000 unique authors. The number of researchers producing knowledge for the region has increased steadily throughout the years with numbers reaching close to 8,000 authors in 2013 compared to less than 2,000 authors in 2004. Interestingly, ASEAN’s plant biotechnology publications have mostly been published in open source journals such as Plos One. ASEAN’s plant biotechnology publications have been cited more than 117,000 times with the highest citation count observed in 2007. The average citation per publication for plant biotechnology publications of ASEAN (19.81) is more than twice higher than the average CPP of all ASEAN publications (8.4) indicating higher influence of plant biotechnology publications than publications in other research areas.

Country output and ASEAN research investments

We then classified the 10 ASEAN countries into three groups based on expenditures on research and development (R&D) (UNESCO Institute for Statistics, 2015): (1) high income countries (HIC) with R&D spending more than 1% of gross domestic product (GDP); (2) middle income countries (MIC) with R&D spending of 0.1 to 0.9% of GDP; and (2) lower middle-income countries (LMIC) with R&D spending of 0.0 to 0.09% of GDP. A significant difference on the publication output in plant biotechnology of HICs with larger R&D investments was noted compared with that of LMICs with less research investments (Table 1). Thailand produced the most number of publications ($n = 2489$). Malaysia and Singapore are the other top three ASEAN producers with more than 150 PPY and CAGR of 29% and 9%, respectively. Philippines with a CAGR of 8% and Vietnam with a CAGR of 19% produced an average of 75 and 41 PPY, respectively. LMICs, namely Brunei Darussalam, Cambodia, Laos, and Myanmar experienced no growth during the ten-year period and have only produced an average of 1-2 papers per year. Interestingly, Indonesia despite its low R&D investments, hence, classified as a LMIC here, was able to produce 61 PPY and is growing at 12% CAGR. The number of authors contributing to ASEAN publications except the LMICs namely: Brunei Darussalam, Cambodia, and Laos, is growing. An increase in the number of contributing authors was especially noted for Malaysia; the country’s number of authors from 2004 to 2013 has increased almost 15 fold.

HICs with higher number of publications received more total citations than lower income countries. Singapore is the most highly cited in plant biotechnology followed by Thailand, Malaysia, and Philippines. With the exception of Indonesia, other LMICs received the least amount of citations for their plant biotechnology publications during the last two decades.

Table 1. Comparison of 2004 and 2013 article output, CAGR, and citation count for ASEAN.

Country	Country classification	Publication output	2004	2013	CAGR	No. of authors	Citation count
Malaysia	MIC	2,199	39	510	29%	10,511	14,584
Vietnam	MIC	418	14	83	19%	2,474	3,957
Thailand	MIC	2,489	108	377	13%	12,688	27,863
Indonesia	LMIC	611	33	104	12%	3,421	7,208
Myanmar	LMIC	23	1	3	12%	100	180
Singapore	HIC	1,594	101	234	9%	10,953	49,094
Philippines	MIC	757	46	104	8%	4,444	14,492
Cambodia	LMIC	6	1	0	-100%	64	135
Brunei	LMIC	35	0	0		30	157
Laos	LMIC	10	0	3		136	186
Total		7,907					117,856

Note: CAGR of Cambodia and Brunei resulted in undefined values and left blank in this table. Source: Scopus

The topmost institution publishing plant biotechnology-related articles in the region are mostly local public research universities (e.g. University Brunei (Brunei), Bogor Agricultural University (Indonesia), National University of Laos (Laos), University of Malaya (Malaysia), Yezin Agricultural University (Myanmar), National University of Singapore (Singapore), and Mahidol University (Thailand). For Cambodia, Vietnam and Philippines, the top producers of publications on plant biotechnology were research institutions and include Cambodian Agricultural Research and Development Institute, Institute of Biotechnology, and International Rice Research Institute (IRRI). The two former institutions are national leading research institutions in bioscience and plant biotechnology while IRRI is an international research organization.

Collaboration

Guided by a decision tree adapted from Lan (2014), we distinguished four types of research collaboration: (1) domestic - in which all authors are in the same country; (2) regional – in which one ASEAN author co-authored with another ASEAN country; and (3) international – in which authors in the ASEAN countries published together with at least one author from another country besides the ASEAN countries. Single authorship and publications that involved intra-institutional co-authorship are not classified as collaboration in this research.

Single author publications and publications that involved intra-institutional co-authorship for ASEAN is very limited; they only constitute 15% of ASEAN's total publications in plant biotechnology. Eighty five percent of ASEAN's total publications in plant biotechnology, on the other hand, involved research collaboration, growing at a CAGR of 15%. Interestingly, the most active institutions that engaged in collaborations in ASEAN are the public universities and institutions of higher education; these institutions have also been noted earlier to be publishing most and the active generators of knowledge for ASEAN. These results confirm observation that plant biotechnology research in ASEAN countries is increasingly conducted now by a group of collaborating researchers rather than by a single researcher (Katz & Martin, 1997; Glänzel, 2001).

The region's co-authored publications that involved domestic partnership are growing at a CAGR of 15%. Six ASEAN members were engaged in domestic collaborations with

Malaysia, Thailand, and Singapore having the highest % shares of domestic collaborations at 42%, 37%, and 20%, respectively. Brunei Darussalam, Cambodia, Laos, and Myanmar have no record of domestic collaborations.

ASEAN publications that involved regional collaboration are very limited with less than 1% of the total collaborations of ASEAN. The highest number of publications that involved regional collaborations was recorded in 2013 ($n = 21$); there was no regional collaboration noted for 2007 and 2008. Ironically, 2007-2008 were the early years of the adoption of ASEAN's Economic Blueprint, which serve as the guide for the establishment of the ASEAN Economic Community. All the higher income countries have co-authored with another ASEAN country although numbers are quite limited (Figure 1). Philippines and Thailand have collaborated mostly with all of the ASEAN countries except Brunei Darussalam. Laos and Myanmar are two of the most active in regional collaborations despite their late membership to the regional association. Both countries have strong regional collaborations with Thailand, their closest ASEAN neighbor; Laos and Thailand used to belong to one country (Siam) and have basically the same language. Brunei has no record of collaborations with any of the ASEAN members.

The region has a very high rate of international collaboration in plant biotechnology research during 2004-2013 at 65% and the rate of collaboration is growing at a CAGR of 11 %. Similar with domestic and regional collaborations, the highest number of publications that involved international collaborations was recorded in 2013 ($n = 227$) while the least was recorded in 2004 ($n = 717$). ASEAN has partnered with 115 countries that are in varying stages of economic development. U.S. remains to be the main international research partner of choice among ASEAN countries. ASEAN is also tapping into the research expertise and resources of other Asian nations like Japan, China, South Korea, and India and advanced countries like United Kingdom, France, Germany, Canada, and The Netherlands. Arunachalam and Doss (2000) had the same observation and stated that Asian countries are fast increasing their share of worldwide international collaboration in science and expanding its collaboration beyond the traditional collaboration with advanced nations such as the United States.

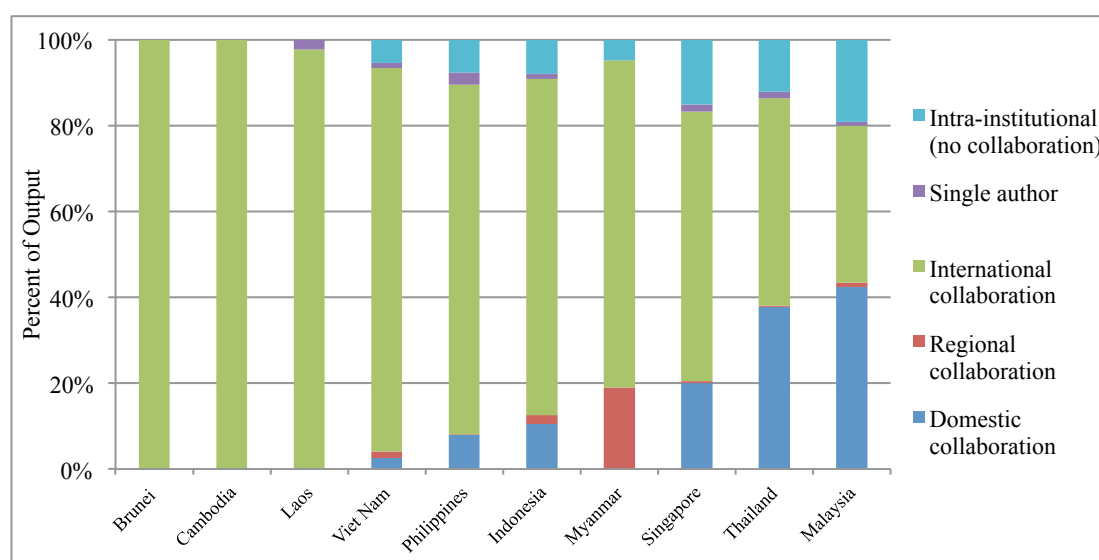


Figure 1. Percentage of different types of collaboration for individual ASEAN countries in plant biotechnology, 2004-2013. Source: Scopus

Brunei Darussalam, Cambodia and Laos are particularly noted for very high international collaboration. There are many justifications for this high collaboration rate and may include

the need for complementary and synergistic research expertise, greater visibility in the international plant biotechnology arena, and greater research output despite limited research investments. Interestingly, the higher income countries and the top ASEAN producers, namely Malaysia, Thailand, and Singapore have lower scientific output with the international community compared with other ASEAN countries, which validates observation that these countries have now higher domestic research capability, hence, would not need as much international collaboration as lower income countries. As expected, ASEAN publications that involved international partnerships received the highest citation count ($n = 86,423$) supporting earlier research while publications that involved regional collaborations received the least citation count ($n = 547$). It is interesting to note that despite the regional collaborations involving more authors and one or more ASEAN countries, the citation count was lower compared to single authored publications. This can indicate the less quality and influence of publications resulting from regional partnerships.

Conclusion and Recommendations

Using bibliometric data for the period 2004-2013 sourced from the research abstract database, Scopus, and deconstructed through a non-commercial home-grown publication parser tool, this paper investigates ASEAN's research output, influence and research collaboration in the area of plant biotechnology. Analysis of the 10-year period indicated an increase in ASEAN plant biotechnology-related scientific output. The publication activity obviously varies from country to country but evident that it is linked with R&D investments: higher income countries such as Singapore produced more publication than lower middle-income countries such as Brunei Darussalam. Most of the knowledge producers of ASEAN were from local research institutions, which are a good indication of improvements in domestic research capability and increase knowledge generation activity among this group. The relatively stable trend of publication generation and increasing R&D investments in countries such as Singapore, Thailand and Malaysia, likewise, provides a good indication that more research output can be expected from these countries. The growth of the publication records especially of Indonesia and Vietnam supports the increasing commitment of these countries and their researchers to contribute in advancing the plant biotechnology field. Philippines need to push and incentivize its local research and academic institutions to produce more and increase their scientific output and not rely on international institution to boost the country's scientific productivity. Brunei Darussalam, Cambodia, Laos, and Myanmar need to improve their research infrastructure and level up their research investments to catch up with other ASEAN countries.

The increasing number of collaborative research teams and number of contributing authors based on co-authorship data in ASEAN publications over the course of the 10-year period, however, is an encouraging result. It represents an increase in the pool of researchers and a change in the balance of research focused more on collaborative research teams among ASEAN researchers and their partners and not on lone scientist.

All the 10 ASEAN countries are actively engaged in research collaboration in plant biotechnology although in varying degrees. The publication output by countries in terms of the collaboration types: domestic, regional and international, differ and is also noted to be linked with status of economic development. Domestic collaborations are very strong for higher income countries with higher R&D investments while lower income countries with lower research investments tend to publish more with their international counterparts. There is more preference for collaboration with more advanced nations but at least the region has expanded its collaboration beyond the United States.

Regional partnerships are, however, very limited, and can be a concern for ASEAN's goal of integration. ASEAN regional collaboration still lag behind in terms of productivity and

quality research in plant biotechnology, which is very evident from the region's low research output and citation count for publications co-authored among ASEAN researchers. Higher regional collaboration rate is only observed to countries that are in close proximity to each other, with common language, and with historical links. Kumar, Rohani, & Ratnavelu (2014) found the same scenario after doing bibliometric work in the field of economics. The low regional collaboration was also mentioned in one of the latest reports by the Asian Development Bank, *Regional Cooperation and Cross-Border Collaboration in Higher Education in Asia: Ensuring that Everyone Wins* (Asian Development Bank, 2012). Hence, it remains to be seen whether regional collaboration will serve as an important platform for continuing to modernize plant science in ASEAN and sharing knowledge in plant biotechnology. More investments in research cooperation, funding mechanisms for regional plant biotechnology research, and other regional incentives need to be setup so ASEAN can realize the goal of its regionalization agenda. Regular quantitative monitoring of inputs and outcomes of research in ASEAN is likewise encouraged to monitor research performance and help in developing research management and science policies, particularly in economic development. Additional research focused on mapping of research collaboration network among ASEAN researchers and their global partners, and a brain circulation study can be done to understand the mobility of ASEAN researchers and whether such movement helps in increasing regional productivity and collaborations and whether such benefits flow back to ASEAN. Furthermore, a qualitative study that would determine other factors that influence an ASEAN researcher to collaborate with another ASEAN researcher or a global partner is suggested.

References

- Arunachalam, S. & Doss, M.J. (2000). Mapping international collaboration in science through coauthorship analysis. *Current Science*, 79(5), 621-628.
- Asian Development Bank. (2012). *Regional Cooperation and Cross-Border Collaboration in Higher Education in Asia: Ensuring that Everyone Wins*. Retrieved June 15, 2015 from: <http://www.adb.org/sites/default/files/publication/29931/regional-cooperation-higher-education-asia.pdf>
- Elsevier. (2014, December 26). *Scopus*. Retrieved June 15, 2015 from: <http://www.scopus.com/>
- Erbisch, F. & Maredia, K. (1998). *Intellectual Property Rights in Agricultural Biotechnology*. Wallingford: CAB International.
- Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics*, 51(1), 69-115.
- Hautea, R. & Escaler, M. (2004). Plant biotechnology in Asia. *AgbioForum*, 7(1 and 2), 2-8.
- Katz, J. & Martin, B. (1997). What is research collaboration? *Research Policy*, 26(1), 1-18.
- Kumar, S., Rohani, V.A., & Ratnavelu, K. (2014). International research collaborations of ASEAN Nations in economics, 1979–2010. *Scientometrics*, 101(1), 847-867.
- Moed, H., Glänzel, W., & Schmoch, U. (2004). *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems*. New York, Boston; London, Moscow: Kluwer Academic Publishers.
- National Agricultural Library. (2014, December 22). *National Agricultural Library Agricultural Thesaurus Library*. Retrieved June 15, 2015 from: <http://agclass.nal.usda.gov/agt.html>
- NIFA. (2014, December 6). *Glossary of Biotechnology Terms: NIFA*. Retrieved June 15, 2015 from: http://www.csrees.usda.gov/nea/biotech/res/biotechnology_res_glossary.html
- Sutton, T.C. (2013). Publication Parser Tool for Scopus. Version 2.0, Pullman, WA.

Science and Technology Indicators In & For the Peripheries. A Research Agenda

Ismael Rafols^{1,2}, Jordi Molas-Gallart¹ and Richard Woolley¹

i.rafols@ingenio.upv.es

¹Ingenio (CSIC-UPV), Universitat Politècnica de València, València (Spain)

²SPRU (Science and Technology Policy Research), University of Sussex, Brighton (UK)

jormoga@ingenio.upv.es, ricwoo@ingenio.upv.es

¹Ingenio (CSIC-UPV), Universitat Politècnica de València, València (Spain)

Abstract

This paper aims to propose a research agenda that explores the problems that emerge when S&T indicators are used in peripheral contexts, that is, in geographical or social spaces that are somehow marginal to the centres of scientific activity. In these situations evaluators and decision-makers are likely to use indicators that were designed to reflect variables relevant in the dominant social and geographical contexts --i.e. in the leading countries, languages, disciplines, etc.--, but that are usually not adequate in peripheral contexts. We propose to examine various dimensions of periphery. First, the cognitive dimension: areas of research, such as the humanities that capture less attention (and resources) than the more prestigious disciplines, such as molecular biology. Second, the geographical dimension: e.g. global south vs. global north, regions vs. metropolises. Third, the social group dimension: women, the poor, or perhaps the elderly have social needs that are different from those of richer or more powerful groups --and the problems affecting the former tend to be less researched than those of the latter. The research agenda proposed would investigate the mechanisms by which performance indicators tend to be biased against the peripheries (e.g. bias in language, journal or topic coverage in conventional databases). We suggest how these biases may suppress scientific diversity and shift research towards a higher degree of homogeneity.

Conference Topic

Science policy and research assessment

Introduction

Science and technology indicators are becoming increasingly used over a wide variety of contexts as research activities become prominent in a larger range of countries, a broader set of organisations, and over a wider range of disciplines or topics (Sa, Kretz et al., 2013). Given that the indicators used in new contexts are often the same, or close adaptations of the indicators used in the traditional disciplines, elite universities and dominant scientific countries, one may wonder about their validity (i.e. adequacy of the indicator to the concept/object is supposed to measure) and their robustness (or sensitivity to contingency in the measuring conditions) (Gingras, 2014).

In this work-in-progress contribution, we propose that many of the new contexts where indicators are used constitute what we call the peripheries or the margin of the research system: spaces that have less visibility, less prestige and/or less resources. As peripheries, these spaces have not had the capacity or influence to develop home-grown indicators suited for their activities -- and are instead relying on indicators borrowed from the central or dominant disciplines and/or countries. For example, it is a recurrent debate in policy to which extent scientometric indicators can be used in the social sciences and humanities (Martin, Tang and Morgan, 2010). Another recurrent example is the case of peripheral countries such as Brazil, where studies have showed that publication practices and citations differed significantly from those in the leading scientific nations, given that they "are significantly influenced by factors "external" to the scientific realm and, thus, reflect neither simply the

quality, influence nor even the impact of the research work referred to." (Velho, 1986, p. 71; see also Velho & Krigge, 1984).

In this contribution we explore dimensions in which the use of indicators in peripheral contexts may be problematic, providing misleading information for research assessment or strategy development. In these contexts, we propose that alternative methods should be explored and potentially developed to create new indicators that are fit for purpose.

This exploration will be developed into the central research agenda for a joint conference of the networks RICYT (the Ibero-American network of Science and Technology Indicators, <http://www.ricyt.org>) and ENID (the European Network of Indicators Designers, <http://enid-europe.org>) to be celebrated in Valencia between 14 to 16 December 2014.

A relational and multidimensional conceptualisation of periphery

The Oxford English Dictionary defines "periphery" as

"The region, space, or area surrounding something; a fringe, margin. Now chiefly: the outlying areas of a region, most distant from or least influenced by some political, cultural, or economic centre."

Its cousin, the Oxford Dictionary of English provides a slightly different definition:

"A marginal or secondary position in, or aspect of, a group, subject, or sphere of activity."

There is already a long history of grappling with the question of peripheries in relation to global social and economic change and development (Prebisch, 1949). Science studies in Latin America have long discussed their peripheral situation and how it meant that their scientific knowledge was dependent, "transplanted" and thus often not properly adapted to their domestic needs -- rather the needs of the Northern countries exploiting their economic resources. For example Vessuri (2004, p. 174) explains that:

"Irrespective of their capabilities, these scientific thinkers were "peripheral" in three senses: in their marginal position in the outer ridges of European culture; in their partial commitment to the scientific endeavour (forced by the immediate pressures for survival in the middle of often unstable contexts, and the economic and political urgencies of new nations); and in their role as agents for the exploitation of natural resources of economic interest for the European centres of power, who gave them legitimacy and support." (Our translation from Spanish)

A noticeable characteristic of this description is the multidimensional nature of the "sense" or spaces of the peripheries of Latin American scientists: culturally (or cognitively), institutionally (partial commitment), in economic terms (unstable resources and dependent on European funding) and in the topics addressed (those of interest to the centres of power).

These definitions suggest two important traits of the notion of periphery, as illustrated by Vessuri's quote above. First, it refers to a situation that is somehow marginal, far from the centre, and where, consequently, less attention is paid. The periphery is therefore *always defined in relation to a centre* where the main locus of the relevant activity resides.

Second, the concept can *relate to many different dimensions* (political, cultural, economic, different "spheres of activity"). In turn these dimensions may or may not be linked with a geographic location; for instance a centre of economic activity will be a specific geographic location. Geographic locations tend to be centre (or periphery) for a variety of dimensions: it is common for political, economic and cultural activities to cluster around geographical centres of power and influence. Similarly, peripheral regions will be peripheral along several dimensions and so the application of the term peripheral to a region has come to indicate a situation of structural disadvantage with broad economic, political and social implications. Developing countries were long ago described as "the" periphery, but within every geographical region we can also encounter peripheral zones (Southern European and Eastern European countries as peripheral to the European Union, or relatively poor regions as

peripheral within their country). Yet, not all dimensions will be correlated for a specific locality. Cambridge is a geographic centre of learning and research (a centre in a cognitive dimension) but, as a city, it is not a centre of political power, although the social group of Cambridge alumni, lecturers and researchers are part of both a political and a cognitive centre. Also, not all relevant dimensions need to have a geographical expression. One can think for instance of social dimensions like gender or class that can be interpreted under the lenses of centre and periphery but are not associated with specific geographic localities. We can therefore refer to peripheral social groups (the disenfranchised, the poor...) whose economic and social needs will be different from those of richer or more powerful communities, even when part of this groups may be located in centres of political power (e.g. the poor neighbourhoods in Washington DC).

Similarly, cognitive dimensions are not necessarily associated with geographic locations; for instance, cognitive peripheries would include areas of research that do not capture the attention of mainstream politicians and receive more limited resources. From this perspective, many fields in the humanities could be considered a peripheral field of knowledge when compared to mainstream natural or engineering sciences.

How conventional indicators are problematic in the peripheries

As we have seen, the notion of a periphery is thus fundamentally a relational one. A periphery is always constituted in relation to a centre, or core. From an indicator perspective, the same entity may thus be peripheral or central depending on the frame of analysis. A particular region may be the centre of nanomaterials research in a particular country, but peripheral in relation to global nanomaterials research, for example. Whether the region is depicted as periphery or centre depends on the frame of comparison. A problem with the *use of indicators* is thus the risk of inappropriate comparisons that can render important activities as relatively trivial.

A second problem relates to whether what is being measured about a particular entity is relevant knowledge in terms of the needs, objectives or valued activities of that entity. The application of an indicator constructed to reflect the needs, objectives or valued activities of another entity may not produce useful information – only a mismatched comparison. A problem with the *content of indicators* is thus the risk of inappropriate comparisons that can render important activities as relatively invisible or lacking in impact. The use of indicators can thus play a role in *constituting peripheries*.

Our goal in this section is to analyse how indicators developed to assess policies and activities related to Science and Technology address peripheral spaces and whether they have constitutive (intended or unintended) effects on these peripheries. We therefore need to identify the dimensions that are relevant to the conduct of S&T.

Each periphery faces its own knowledge generation and application context and may be better analysed using specific, tailored indicators. Yet, by and large they need to rely on indicators, and analytical models developed for the studies of "centre" spaces. Evaluators and decision-makers are likely to use indicators that were designed to reflect variables relevant in the dominant social and geographical contexts --i.e. in core regions, languages, disciplines, etc.--, but that are usually not adequate in peripheral spaces.

Let us see some examples of dimensions where use of indicators in the periphery is problematic.

Language

Language has long been known to be a major problem for performance measures, given that non-English articles tend to be much less cited. Van Leeuwen et al. (2001) showed that the inclusion or not of non-English publications in the analysis of citation impact has a major influence in the outcomes of indicators. Van Raan et al. (2011) showed that this also had

major effects in university rankings. Vasconcelos et al. (2008) showed that language proficiency is highly correlated with citation impact and h-index of researchers. This means that for the purposes of comparison, non-English publication should be excluded in most analysis.

Gender

In many fields of science, women tend to publish less and accrue less citations than men. However, various studies have consistently found that women tend to do more interdisciplinary research (e.g. Leahey, 2007; Van Rijnsvoort and Hessels, 2011). Hence, the effect of gender on performance depends on the indicators choice: if publications and citations are taken as a measure of the value of a contribution, the indicators will tend to disadvantage female researchers.

Basic vs. applied vs. research

Applied studies tend to cite fundamental studies more than the reverse. As a result, fundamental research tends to appear as more central in global science maps (Rafols, Porter and Leydesdorff, 2010). This is possibly a perception bias without serious repercussions. The serious problem is that even within a given scientific field as defined by conventional classifications such as Web of Science Categories, applied research tends to be significantly less cited than fundamental research (van Eck et al., 2013).

Interdisciplinary research

Interdisciplinary research can be thought of as peripheral to the extent that it is published in areas outside the disciplinary cores. It turns out that interdisciplinary research tends to be published in journals with lower rating in journal rankings and, within a field, with journals with a lower Journal Impact Factor (Rafols et al., 2012). As a result interdisciplinary research tends to be in a disadvantage when using this type of journal-based indicators (with citation indicators, the effect may vary as it depends on relative citation rates between fields that are being cross-fertilised).

Conclusions

S&T indicators tend to be biased against organisations, countries or disciplines in the periphery. This is possibly due to the fact that indicators were not initially designed for the peripheries. At the same time, the use of these indicators in assessments linked to the distribution of resources can have constitutive effects, reinforcing for instance the peripheral character of a region or discipline. These remain unresolved problems for S&T indicators and their use in evaluation. In this contribution we shed light on this bias in multiple dimensions, in order to foster critical awareness of the problems caused by biases as well as the development of context sensitive indicators (Lepori & Reale, 2012).

Acknowledgments

Ismael Rafols thanks Diego Chavarro (who is doing a thesis on the Journal Indexing Systems in Latin America) for many illuminating discussions on indicators in developing countries.

References

- Gingras, Y. (2014). Criteria for evaluating indicators. In B. Cronin & C. Sugimoto (Eds.), *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact* (pp. 109–126). Cambridge, MA and London, UK: MIT Press.
- Martin, B. R., Tang, P., Morgan, M. et al. (2010). *Towards a Bibliometric Database for the Social Sciences and Humanities – A European Scoping Project* (A report for DFG, ESRC, AHRC, NWO, ANR and ESF). Brighton, UK: SPRU.

- Leahey, E. (2007). Not by Productivity Alone: How Visibility and Specialization Contribute to Academic Earnings. *American Sociological Review*, 72(533-561).
- Lepori, B. & E. Reale (2012). S&T indicators as a tool for formative evaluation of research programs. *Evaluation* 18(4): 451-465.
- Prebisch, R. (1949). Interpretación del proceso de desarrollo latinoamericano. *CEPAL*, Santiago de Chile, Chile.
- Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., & Stirling, A. (2012). How journal rankings can suppress interdisciplinarity. The case of innovation studies and business and management. *Research Policy*, 41(7), 1262–1282.
- Sá, C. M., A. Kretz, et al. (2013). "Accountability, performance assessment, and evaluation: Policy pressures and responses from research councils." *Research Evaluation* 22(2): 105-117.
- Van Eck, N. J., Waltman, L., van Raan, A. F. J., Klautz, R. J. M., & Peul, W. C. (2013). Citation Analysis May Severely Underestimate the Impact of Clinical Research as Compared to Basic Research. *PLoS ONE*, 8(4), e62395. doi:10.1371/journal.pone.0062395
- 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.
- Van Raan, A. F., Van Leeuwen, T. N., & Visser, M. S. (2011). Severe language effect in university rankings: particularly Germany and France are wronged in citation-based rankings. *Scientometrics*, 88(2), 495-498.
- Van Rijnsoever, F. J., & Hessels, L. K. (2011). Factors associated with disciplinary and interdisciplinary research collaboration. *Research Policy*, 40(3), 463–472.
- Vasconcelos, S. M., Sorenson, M. M., Leta, J., Sant'Ana, M. C., & Batista, P. D. (2008). Researchers' writing competence: a bottleneck in the publication of Latin-American science?. *EMBO reports*, 9(8), 700-702.
- Velho, L., & Krige, J. (1984). Publication and citation practices of Brazilian agricultural scientists. *Social Studies of Science*, 14(1), 45-62.
- Velho, L. (1986). The “meaning” of citation in the context of a scientifically peripheral country. *Scientometrics*, 9(1), 71-89.
- Vessuri, H. (2004). La hibridización del conocimiento. La tecnociencia y los conocimientos locales a la búsqueda del desarrollo sustentable. *Convergencia*, 35(May-August), 171–191.

Patterns of Internationalization and Criteria for Research Assessment in the Social Sciences and Humanities

Gunnar Sivertsen

gunnar.sivertsen@nifu.no

Nordic Institute for Studies in Innovation, Research and Education
(NIFU) P.O. Box 5183 Majorstuen, N-0302 Oslo (Norway)

Abstract

This paper investigates the developments during the last decades in the use of languages, publication types, and publication channels, in the social sciences and humanities (SSH). The purpose of the study is to develop an understanding of the processes of internationalization and to apply this understanding in a critical examination of an often used criterion in research evaluations in the SSH: Coverage in Scopus or Web of Science is seen in itself as an expression of research quality and of internationalization. This extrinsic ‘coverage criterion’ is beyond the control of academia and without support in analysis of how research quality and relevance is achieved through scholarly publishing in the SSH. It needs to be replaced by intrinsic criteria based on the SSH’s own concepts of field-specific research excellence and societal relevance. The study will demonstrate this by using data from scholarly publishing in the SSH that go beyond the coverage in the commercial data sources by giving a more comprehensive representation of the SSH.

Conference Topic

Science policy and research assessment

Introduction

The presence of publications in Scopus or Web of Science (WoS) has increasingly become a criterion in evaluations of research in the social sciences and humanities (SSH). Some countries have even installed protocols for research evaluation or performance-based funding models where publications that are indexed by the commercial databases are treated separately in indicators of “internationalization” and “research quality”. In other countries, there is a general belief that research quality can be promoted in the SSH by expecting more publications in the limited number of international journals that have been selected for indexing. Consequently, for several years already, Elsevier and Thomson Reuters have experienced a pressure from researchers in the SSH to have more journals indexed. Both providers have responded by increasing the coverage of journals and book series, and, recently, even of books in the SSH. However, the coverage of the scholarly publication output in the SSH is still limited (Sivertsen, 2014). The shortage is mainly due to the more heterogeneous scholarly publication patterns in the SSH where publishing in international journals is supplemented by book publishing and the use of journals in the native languages (Hicks, 2004; Archambault et al, 2006; Engels, Ossenblok & Spruyt, 2012; Sivertsen & Larsen, 2012; Sivertsen, 2014).

Just as with the abuse of Journal Impact Factors in research assessment in science, technology and medicine (STM), the ‘coverage criterion’ in the SSH represents an artefact which is external to and beyond the control of the scholarly norms and standards that it is sought to represent. It creates unnecessary tensions between fields in the SSH with different degrees of coverage in the databases. It also creates debates about what will happen to the use of books and native languages in the SSH. In these debates, the general development towards publishing in journals covered by Scopus or Web of Science is often perceived as “inevitable” and driven by new evaluation regimes, not by internal scholarly standards. In this study, I will develop an understanding of the processes of internationalization in the SSH which is

independent of the ‘coverage criterion’ and instead related to concepts of field-specific research excellence and societal relevance in the SSH.

Methods

For the purpose of this study, data are needed that give a complete representation of scholarly publishing in the SSH, also of publications in books, series and journals not covered by Scopus or Web of Science. In 2005, Norway was the first country to establish a national information system with complete quality-assured bibliographic data covering all peer-reviewed scholarly publishing in the total higher education sector (Schneider, 2009; Sivertsen, 2010). This national system, which is now called CRISTIN (Current Research Information System in Norway) and has been expanded beyond the higher education sector, provides the main source of data for this study.

The methodology of the bibliographic data collection in the Norwegian CRISTIN database (www.cristin.no) has been published earlier (Sivertsen, 2010; Sivertsen & Larsen, 2012; Sivertsen, 2014). Scientific and scholarly publications of all fields are covered completely according to an agreed definition. Among other criteria, the definition demands originality and scholarly format in the publication and peer-review in its publication channels. All publication channels (journals, series, book publishers) and publication types (see below) are standardized in the database.

Humanities is defined in our study as the disciplines included in this major area in the OECD Field Classification.¹ The *Social Sciences* are defined in the same way with the exception of Psychology, which we have not included in this study. Note that Law and Educational Research are classified as social sciences by OECD.

Two supplementing data sets (A, B) will be used, each of them for a more specific purpose:

- A. For the analysis of publication patterns in the SSH down to the level of individual researchers, we use data from the above-mentioned CRISTIN system which cover the four years 2010-2013. The unit of analysis is publications per researcher within a variable of three publication types (articles in journals or series with ISSN; articles in books; books) and a dichotomous variable of languages (Norwegian (the native language); International languages). The data include 1,895 unique researchers in the humanities with 7,145 unique publications, and 3,229 unique researchers in the social sciences with 11,817 unique publications.
- B. For the analysis of the development of publication patterns in the SSH over time, we use data that are defined and collected in the same way as in data set A, but aggregated at the level of disciplines. The data cover the years 2005-2011. The unit of analysis is publication per discipline (and major area) with the same variables of publication types and languages as in data set A. Data set B includes 14,558 unique publications in the humanities and 19,450 unique publications in the social sciences.

Results, Part I: Characteristics of the Publication Patterns in the SSH

As seen in *Table 1*, publications in journals and series represent a little more than half of the publications in the humanities and two thirds of the publications in the social sciences, indicating that book publishing is important as well, especially in the form of articles in books (edited volumes). There are, however, just as wide differences *within* each of the two major areas: Only 45 per cent of the publications in History are in journals, compared to 61 per cent in Linguistics. In Sociology, only 46 per cent of the publications are in journals, compared to 75 per cent in Economics.

¹ OECD: REVISED FIELD OF SCIENCE AND TECHNOLOGY (FOS) CLASSIFICATION IN THE FRASCATI MANUAL, version 26-Feb-2007, DSTI/EAS/STP/NESTI (2006)19/FINAL.

Table 1. Number and percentage publications per publication type. Based on data set A.

	<i>Humanities</i> <i>N</i>	<i>Humanities</i> <i>%</i>	<i>Soc Sci</i> <i>N</i>	<i>Soc Sci</i> <i>%</i>
Books	328	4.6 %	273	2.3 %
Articles in books	2,861	40.0 %	3,640	30.8 %
Articles in journals or series	3,956	55.4 %	7,904	66.9 %
Total	7,145	100.0 %	11,817	100.0 %

The scholarly publication types in the SSH are often discussed as if they represent alternatives to each other: Is the use of one of the publication types increasing at the cost of the others? Are monographs becoming obsolete in the SSH? Before we study the trends, we shall observe an indication that the publication types are supplementing each other rather than competing with each other. As seen in Table 2, the numbers and percentages of *the researchers* that actually use a certain publication type are significantly higher than in Table 1, indicating that more than one publication type is often present in the publishing profile of an individual researcher. As an example, although less than a third of the publications in the social sciences are articles in books, more than half of the researchers are using this publication type.

Table 2. Number and percentage of the researchers using a publication type within four years. Based on data set A.

	<i>Humanities</i> <i>N</i>	<i>Humanities</i> <i>%</i>	<i>Soc Sci</i> <i>N</i>	<i>Soc Sci</i> <i>%</i>
Books	297	15.7 %	273	8.5 %
Articles in books	1,187	62.6 %	1,676	51.9 %
Articles in journals or series	1,537	81.1 %	2,775	85.9 %
Total (unique researchers)	1,895		3,229	

Table 3 demonstrates to what degree the publishing profiles of individual researchers include more than one publication type. Even in the social sciences, where journal articles represent two thirds of the output, almost half of the researchers who publish these articles also use other publication types.

Table 3. Number and percentage of the researchers using a publication type that also uses another publication type within four years. The percentages are related to the numbers (N) in Table 2. Based on data set A.

	<i>Humanities</i> <i>N</i>	<i>Humanities</i> <i>%</i>	<i>Soc Sci</i> <i>N</i>	<i>Soc Sci</i> <i>%</i>
Books	265	89.2 %	250	91.6 %
Articles in books	891	75.1 %	1,275	76.1 %
Articles in journals or series	930	60.5 %	1,291	46.5 %

So far, we can conclude that book publishing and journal publishing seem to supplement each other rather than represent alternatives in the SSH. We will return to a possible explanation for this in the discussion at the end.

We now turn to another dimension in the publication patterns of the SSH – the language dimension. In non-English speaking countries, the use of the native language in scholarly

publications is an indication that the publication is mainly oriented at a national or regional audience of readers in which not only peers, but also students, policy makers, professionals, media and a wider public may be reached as well. Since scholarly publications in the native languages are relatively frequent in the SSH, publishing in an international language is, on the other hand, not the normal situation, as in the sciences, but a clear expression of an ambition to reach an international audience of experts in the field.

We proceed as with the publication types and start with an overview of the use of language in publications in Table 4. In both the humanities and the social sciences, the majority of scholarly publications are in the international languages. However, publications in the native language are much more frequent than in the sciences, indicating that such publications have a specific role in the SSH.

Table 4. Number and percentage publications per language type. Based on data set A.

	<i>Humanities</i> <i>N</i>	<i>Humanities</i> <i>%</i>	<i>Soc Sci</i> <i>N</i>	<i>Soc Sci</i> <i>%</i>
International language	4,368	61.1 %	8,666	71.7 %
Norwegian language	2,777	38.9 %	3,418	28.3 %
Total	7,145	100.0 %	11,817	100.0 %

Again, the question may be raised: Are the native and international languages supplementing each other, or are they competing as alternatives? By going down to the level of individual researchers, we can observe in Table 5 that high proportions of the researchers combine both types of languages in their publication practice. While a majority of researchers publish in the international languages, there is *no minority of researchers* publishing in the native language only. Researchers in the SSH are *normally bilingual* in their publication practice (if their native language is not English).

Table 5. Number and percentage of the researchers using international and native languages in their scholarly publications within four years. Based on data set A.

	<i>Humanities</i> <i>N</i>	<i>Humanities</i> <i>%</i>	<i>Soc Sci</i> <i>N</i>	<i>Soc Sci</i> <i>%</i>
International language	1,482	78.2 %	2,687	83.2 %
Norwegian language	1,228	64.8 %	1,725	53.4 %
Total (unique researchers)	1,895		3,229	

A more general conclusion from the results so far, is that although the *majority of publications* in the SSH are published in journals and in international languages, *the majority of researchers* are publishing in books and in the native language as well. Is this picture changing?

Results, Part II: Developments in the Publication Patterns in the SSH

To study the developments, we use data set B, by which it is possible to cover a longer period of time. The general picture is that the publication patterns in the SSH are quite stable, both with regard to publication types (Figure 1) and the use of international versus native languages (Figure 2). In relative shares, the uses of international languages and of journals are increasing, but not by a high rate. In absolute numbers, there is no in reduction book publishing or the use of the native language, since in data set B, which we are using here, there was an increase in the total number of publications by more than 50 per cent between 2005 and 2011.

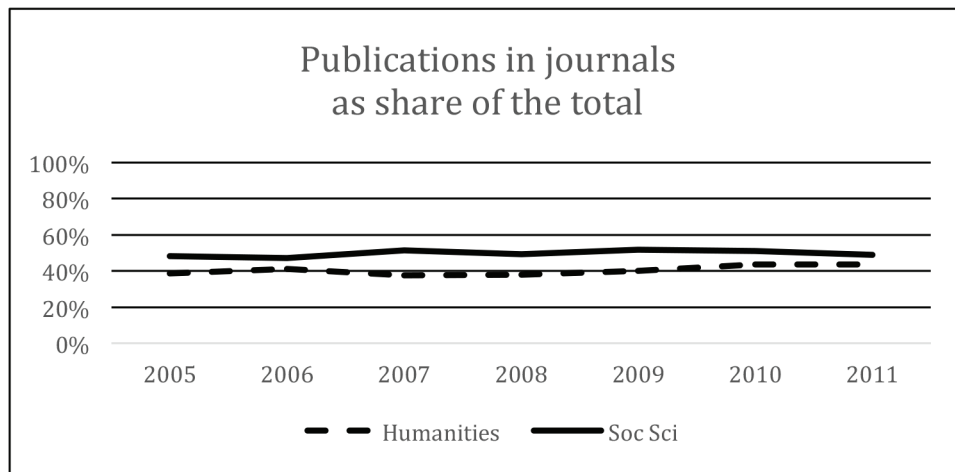


Figure 1. Scholarly publications in journals as a percentage of the total, which also includes articles in books and books. Based on data set B.

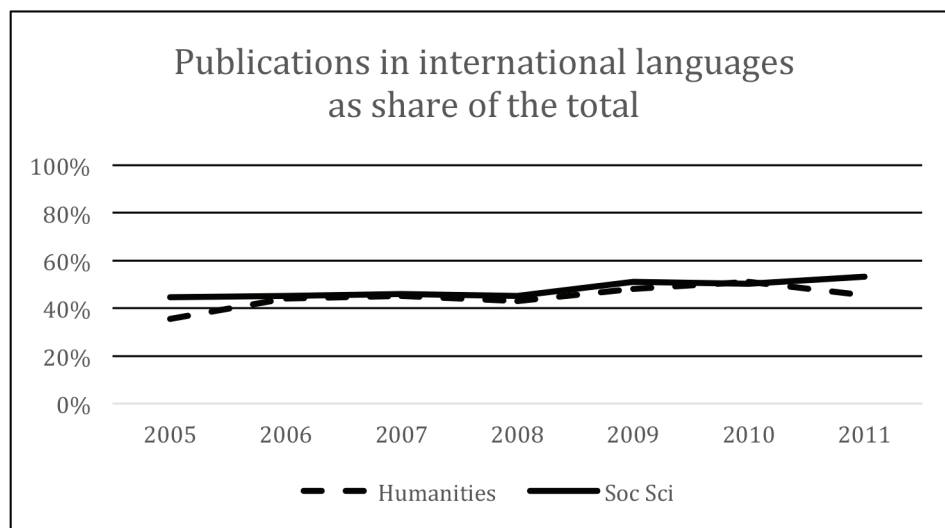


Figure 2. Scholarly publications in international languages as a percentage of the total, which also includes publications in the native language. Based on data set B.

Discussion and Conclusions

The normal publication practice in the SSH, in which both types of languages, and books as well as journals, are used for scholarly publishing by the majority of researchers, seems to prevail during a period of internationalization. The stability of the publication patterns, as well as their differences *within* the SSH (Sivertsen & Larsen, 2012; Ossenblok, Engels & Sivertsen, 2012), indicate that the choice of language and publication type is not just a question of new trends versus old traditions. Publication patterns are more deeply rooted in scholarly norms, methods and practices. The monograph, the edited book and the journal article represent different methodologies that may all need to be used at different times. The choice of language depends on the international scholarly relevance of the research versus the societal relevance for the culture and society being studied. One and the same research project may well contribute with different parts to both dimensions. The SSH would lose their *raison d'être* and societal impact by disconnecting from the surrounding culture and society and mainly communicating in international journals that are only read by peers abroad. At the same time, publishing in those specialized journals on the international level is necessary in

order to be confronted with and inspired by the scholarly standards, critical discussions and new developments among other experts in the field.

In the context of criteria for research evaluation in the SSH, there is a need to accept that none of the alternatives in the two dimensions of the scholarly publication patterns that have been described here – language and publication type - can be regarded as more valuable alternatives. All of them contribute – with different roles and connected to different methodologies, audiences and feedbacks – to research excellence and societal relevance of the SSH. The coverage in Scopus or the Web of Science of the scholarly publishing pattern in the SSH is far from complete (Sivertsen, 2014). Hence, *coverage in a commercial indexing service* should not be used as a criterion for research quality or an indicator of internationalization in the SSH.

References

- Archambault, E., Vignola-Gagne, E., Cote, G., Lariviere, V., & Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics*, 68(3), 329-342.
- Engels, T.C.E., Ossenblok, T.L.B., & Spruyt, E.H.J. (2012). Changing publication patterns in the social sciences and humanities 2000-2009. *Scientometrics*, 93(2), 373-390.
- Hicks, D. (2004). The four literatures of social science. In Moed, H., Glänzel, W., & Schmoch, U. (Eds.) *Handbook of Quantitative Science and Technology Research*. Kluwer Academic Publishers.
- Luukkonen, T., Persson, O., & Sivertsen, G. (1992). Understanding patterns of international scientific collaboration, *Science, Technology & Human Values*, (17), 101-126.
- Ossenblok, T.L., Engels, T.C., & Sivertsen, G. (2012). The representation of the social sciences and humanities in the Web of Science – a comparison of publication patterns and incentive structures in Flanders and Norway (2005–9). *Research Evaluation*, 21(4), 280-290.
- Schneider, J.W. (2009). An Outline of the Bibliometric Indicator used for Performance-based Funding of Research Institutions in Norway, *European Political Science*, 8, 364–78.
- Sivertsen, G. (2010). A performance indicator based on complete data for the scientific publication output at research institutions, *ISSI Newsletter*, 6, 22–8.
- Sivertsen, G. & Larsen, B. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: An empirical analysis of the potential. *Scientometrics*, 91(2), 567-575.
- Sivertsen, G. (2014). Scholarly publication patterns in the social sciences and humanities and their coverage in Scopus and Web of Science. In *Proceedings of the Science and Technology Indicators Conference 2014 Leiden*, ed. Ed Noyons, 598-604. Leiden: Centre for Science and Technology Studies.

Looking for a Better Shape: Societal Demand and Scientific Research Supply on Obesity

Lorenzo Cassi¹, Ismael Rafols², Pierre Sautier³ and Elisabeth de Turckheim⁴

¹ *lorenzo.cassi@uni-paris1.fr*

Observatoire des Sciences et Techniques (HCERES-OST) and CES University of Paris 1 Pantheon-Sorbonne,
Paris (France)

² *i.rafols@ingenio.upv.es*

Ingenio (CSIC-UPV), Universitat Politècnica de València, València (Spain),
SPRU (Science and Technology Policy Research), University of Sussex, Brighton (UK),
and Observatoire des Sciences et Techniques (HCERES-OST), Paris (France)

³ *pierre.sautier@obs-ost.fr*

Observatoire des Sciences et Techniques (HCERES-OST), Paris (France), and Ingenio (CSIC-UPV), Universitat
Politècnica de València, València (Spain)

⁴ *elisabeth.deturckheim@obs-ost.fr*

Observatoire des Sciences et Techniques (HCERES-OST), Paris (France), and INRA, Délégation à l'évaluation,
Paris (France)

Abstract

As science policy shifts towards an increasing emphasis in societal problems or grand challenges, new scientometric tools are required to inform decision-makers. However, while traditional bibliometrics could focus on the knowledge production side (the science supply), grand challenges also demand to investigate the articulation of societal needs. In this paper, we present an exploratory investigation of the grand challenge of obesity -an emerging health problem with enormous social costs. We illustrate a potential approach, showing: (a) how scientific publication can be used to describe existing science supply by using topic modelling based on publication abstracts; (b) how question records in the French parliaments can be used as an instance of social demand; and (c) how the comparison between the two may show (mis)alignments between societal concerns and scientific outputs.

Conference Topic

Science policy and research assessment

Introduction

Tackling complex global problems or grand challenges – such as climate change, food security, poverty reduction, risk of global pandemics – requires not only to increase R&D expenditure, but also the exploration and eventually the coordination of a variety of stakeholders with different areas of expertise and pursuing diverse research avenues. Typically these challenges benefit from the understanding of the physical and biological phenomena underlying a challenge (e.g. the virus and its genes), but also demand an understanding of the environmental and social contexts in which they occur, and the policy networks and instruments available in those contexts (Ely, Van Zwanenberg & Stirling, 2014).

Science policy funding schemes for societal problems or grand challenges seek to align science supply with social problems or needs. Although science is conducted in conditions of incomplete knowledge, it is well documented that certain particular research options are much

better aligned to certain outcomes (Sarewitz, 1996, pp. 31–49). It is, for example, very unlikely that astrophysics be useful for improving health care in malaria. Historically, several lines of inquiry in science policy have explored the alignment between research options and social outcomes, namely related to priority-setting and evaluation of research, but also to broader considerations related to the “supply” and “demand” of policy-relevant science. For this reason, a suitable interpretation of the alignment issue should be based on our understanding of the current state of the science (the *supply*) and what is required to achieve social goals (the *demand*) (Sarewitz & Pielke, 2007). The “demand” side must therefore consider not only the plurality of outcomes, but also various ways of articulating specific science or technology -driven pathways for achieving them. This in turn can refer to a process of public deliberation whereby different outcome preferences or divergent underlying values are made explicit by stakeholders. Similarly, the “supply” side is not just about how much “high-risk, high-return” research should be undertaken, but also about what type of outcomes are more or less *likely* to result from a given line of research. In this article, on the one hand, we apply the concept of *research landscape* (Wallace & Rafols, 2014) in order to map the scientific research on obesity.

On the other hand, we symmetrically map one of the interpretations (representations) of social needs (demand) on obesity. The supply-demand schema can be represented as in Figure 1. Here societal demand and scientific supply are not related directly in one single way. Instead they can relate via a variety of interpretation/representations of the “obesity” social needs. These representations shape science policy and affect actions that may reconcile supply and demand.

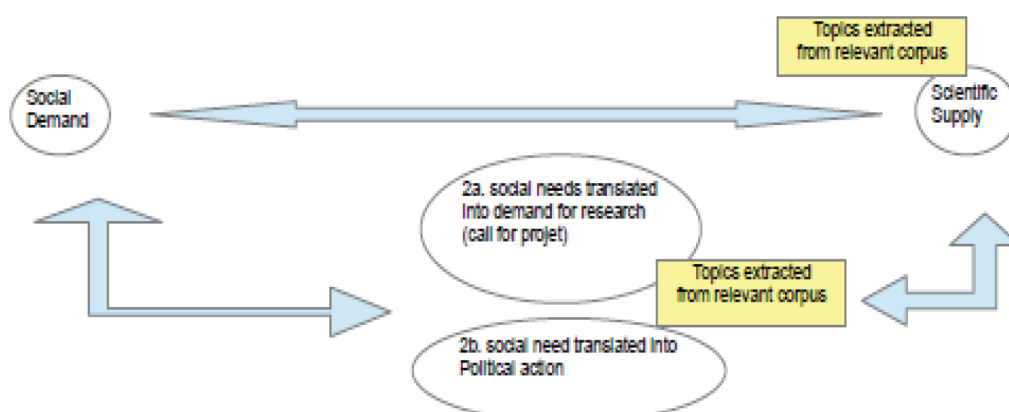


Figure 1. Social demand – scientific supply and political discourse as an example of intermediary representation.

In this paper we investigate the alignment (or lack thereof) between science supply and social demand by mapping, first, the scientific supply via the research landscape of obesity as defined by topic modelling of publication abstracts, and, second, social demand according to political discourse in the French parliaments. These maps of both supply and demand are specific and partial representations used in this preliminary and exploratory study -- other, complementary representations would be possible. For example, supply could be represented by a topic modelling of grants abstracts (as Talley et al., 2011 did for the US National Institutes of Health). And demand could be mapped using newspaper articles, among many other sources.

Data and Methods

Data

In order to define the relevant corpus for obesity, we follow a two-step method. First, we retrieve all publications with indexed MeSH term matching the search *obes** in MEDLINE/PubMed during the 2002-2013 period. This search was performed on October 16, 2014 and it returned 87,315 records.

Then, we launched *medlineR*, a routine based on the R language that allows the user to match data from Medline/PubMed with records indexed in the ISI Web of Science (WoS) database (Rotolo & Leydesdorff, 2015). The routine identified 71,055 WoS records (WoS core collections), with 'article' or 'review' as document types.

Second, we used Leiden's classification system to identify clusters of publications related to obesity. The classification system is constructed at the level of individual publications and clustering is based on direct citations (Waltman & van Eck, 2012) for the period 2000-2013. Obesity publications appear in 4,718 micro-clusters (in which at least one publication is tagged obesity), out of 32,466 micro-clusters for the whole WoS corpus. All the publications from clusters with at least 25% of publications tagged as 'obesity' were considered to be relevant for the study. This threshold of 25% is arbitrary and exploratory. Further explorations will use a lower threshold to test the robustness of this choice. The obesity corpus thus obtained contains 54,424 publications.

Topic modelling

Topic modelling provides a suite of algorithms to discover hidden thematic structure in large collections of texts. A topic model takes a collection of texts as input and it discovers a set of topics (recurring themes that are discussed in the collection) and the degree to which each document exhibits those topics.

Latent Dirichlet Allocation (LDA) is the simplest topic model. The intuition behind LDA is that documents exhibit multiple topics. LDA is a statistical model of document collections that tries to capture this intuition. It is most easily described by its generative process, the assumed random process. A topic is defined as a distribution over a pre-defined vocabulary. Moreover, it is assumed that the topics are specified before data have been generated (technically, the model assumes that the topics are generated first, before the documents). Now for each document in the collection, we generate the words in a two-stage process:

1. Randomly choose a distribution over topics.
2. For each word in the document
 - (a) Randomly choose a topic from the distribution over topics in step #1.
 - (b) Randomly choose a word from the corresponding distribution over the vocabulary.

This statistical model reflects the idea that each document contains multiple topics. Each document exhibits the topics with different proportion (step #1); each word in each document is drawn from one of the topics (step#2b), where the selected topic is chosen from the per-document distribution over topics (step #2a).

The goal of topic modelling is to automatically identify the topics from a collection of documents. The documents themselves are observed, while the topic structure (the topics, per-document topic distributions and the per-document per-word topic assignments) is a hidden structure.

Results on Science Supply

For this study, we fitted a 100-topic model to the 54,424 publications of the obesity corpus. We perform LDA with the R package “topicmodels” and visualize the output using LDavis.

Figure 2 shows a map of these 100 topics. Topics are located close to one another if they are similar in terms of distributions of the words belonging to the selected dictionary. The measure of topic similarity is the matrix of Jensen-Shannon divergences between topics, considered as distributions over words, into two-dimensional coordinates and is represented in a 2d space through multi-dimensional scaling (i.e., principal coordinates analysis).

In addition, a clustering technique is used to cluster topics into research areas. We applied k-means clustering to the topics as a function of their two-dimensional locations in the global topic view with $k=10$. Labels are assigned to clusters. These labels are obtained by extracting the most relevant terms for each cluster of topics, where the term distribution of a cluster of topics is defined as the weighted average of the term distributions of the individual topics in the cluster.

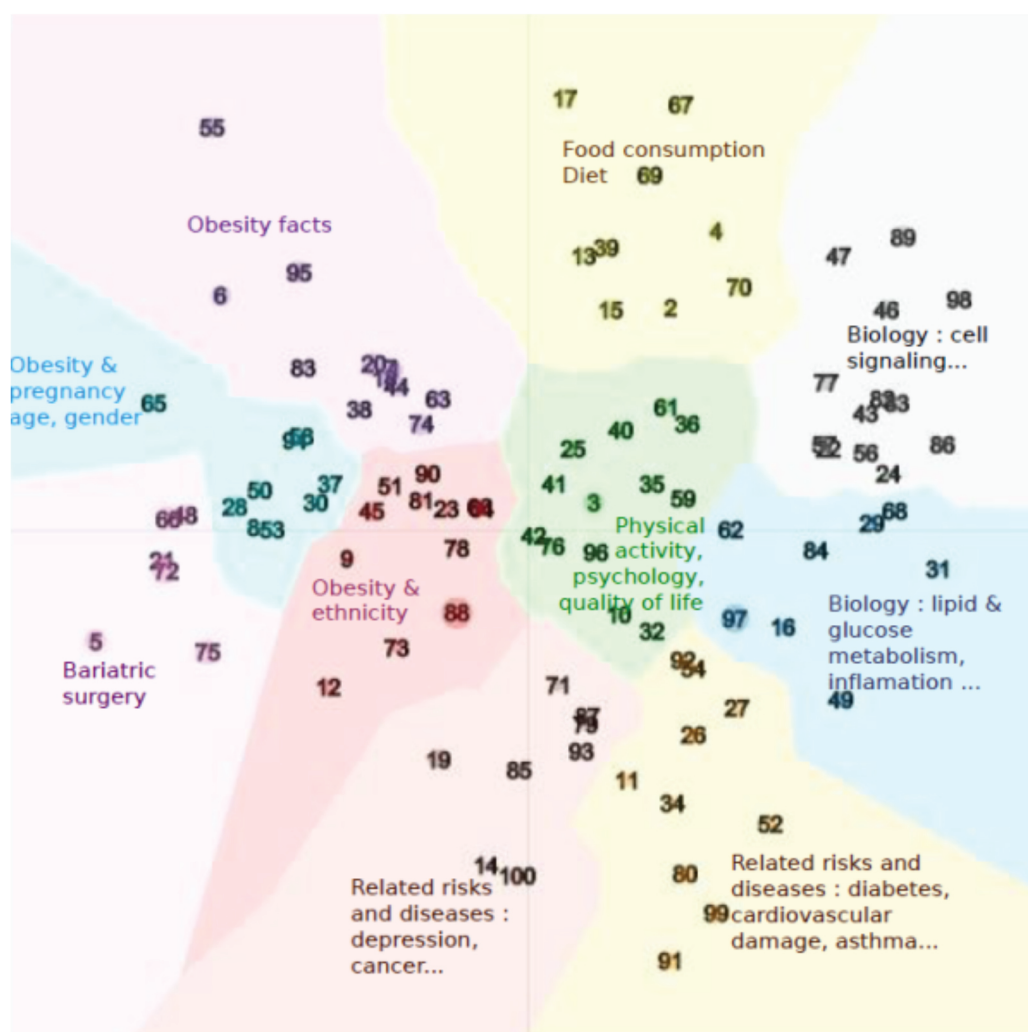


Figure 2. Map of topics of publications on obesity (2003-2013).

Results on the Societal Demand

The same approach has been used to map the social demand. In order to define one possible interpretation, we refer to the questions that the members of the French Parliament (i.e. Assemblée Nationale or Senate) can ask to the government. Deputies and senators publicly question the members of the Government in different ways. The question can be asked during a Parliament session to the government or be written and a Parliament session is not necessary and addressed to one of the ministers. We retrieved this information and built up two datasets.

First, we selected all the questions asked by the Senators where the word *obes** was reported in the public database - with records from 1985 on - which is now available. We got 242 questions from 1992 - year of the first occurrence of 'obesity' in these questions - to 2014. Second, we collected oral and written questions asked by members of the Assemblée Nationale in the last three legislatures, getting: 422 questions (2002-2007), 870 (2007-2012) and 380 (2012 – 2014). The output of the 10-topic model is shown below for the Senate questions.



Figure 3. Map of topics for questions in the French Senate (1992-2014).

Discussion

In the centre of the Figure 2, we have a cluster of topics concerning *Physical activity, psychology and quality of life*, then turning around clockwise we find *Food consumption and diet* and then two clusters concerning mainly topics linked to biology research and further four clusters related to medical and surgery issues. The clusters of topics identified in the research landscape are mainly concerning medical and biological issues and only two clusters seem to deal with social and behavioural determinants of obesity, respectively *Obesity & ethnicity* and *Food consumption and diet*. The political discourse (Figure 3) seems to be organised around topics different from the research landscape. Among the ten topics defined

three main groups are reasonably identified. The first one, on the top part of the graph (i.e., topics number 3, 5, 7 and 10), is concerned mainly with children nutrition and the role of media as in advertising. A second group of topics, on the bottom right of the graph (i.e., topics number 1, 2 and 4), deals with food industry, marketing, and labelling issues. Finally, a third group, at the bottom left (i.e., topics number 6, 8 and 9) is concerned by medical and surgery issues. Only three out of ten topics of political discourse seem to find a counterpart in the research landscape. A preliminary analysis therefore suggests that, while research is concerned about the biophysical mechanisms that lead to obesity, many of the political questions are about the social mechanisms that favour obesity, such as advertisement, beverages, marketing, etc. This may suggest insufficient research regarding the social origin of obesity.

Acknowledgments

We would like to thank Tommaso Ciarli for suggesting to us the use of Parliament database as one of the possible representations of social needs. We thank Ludo Waltman for sharing the article level classification system.

References

- Ely, A., Van Zwanenberg, P., & Stirling, A. (2014). Broadening out and opening up technology assessment: Approaches to enhance international development, co-ordination and democratisation. *Research Policy*, 43(3), 505–518. doi:10.1016/j.respol.2013.09.004
- Rotolo, D., & Leydesdorff, L. (2014). Matching MEDLINE/PubMed data with Web of Science (WoS): A routine in R language. *Journal of the Association for Information Science and Technology* (Forthcoming).
- Sarewitz, D. (1996). *Frontiers of Illusion: Science, Technology and the Politics of Progress*. Philadelphia: Temple University Press.
- Sarewitz, D., & Pielke, R. A. (2007). The neglected heart of science policy: reconciling supply of and demand for science. *Environmental Science & Policy*, 10(1), 5–16.
- Talley, E. M., Newman, D., Mimno, D., Herr, B. W., Wallach, H. M., Burns, G. A. P. C., & McCallum, A. (2011). Database of NIH grants using machine-learned categories and graphical clustering. *Nature Methods*, 8(6), 443–444. doi:10.1038/nmeth.1619
- Wallace, M. L., & Rafols, I. (2014). Research portfolios in science policy: moving from financial returns to societal benefits. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2500396.
- Waltman, L., & van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, 63(12), 2378–2392. doi:10.1002/asi.22748

Does Quantity Make a Difference?

Peter van den Besselaar¹ & Ulf Sandström²

¹*p.a.a.vanden.besselaar@vu.nl*

VU University Amsterdam, Department of Organization Sciences & Network Institute

²*ulf.sandstrom@oru.se*

Royal Institute of Technology, INDEK & Orebro University, Business School, Orebro

Abstract

Do highly productive researchers have significantly higher probability to produce top cited papers? Or does the increased productivity in science only result in a sea of irrelevant papers as a perverse effect of competition and the increased use of indicators for research evaluation and accountability focus? We use a Swedish author disambiguated dataset consisting of 48,000 researchers and their WoS-listed publications during the period of 2008-2011 with citations until 2014 to investigate the relation between productivity and production of highly cited papers. As the analysis shows, quantity does make a difference.

Conference Topic

Indicators; Science policy; Research assessment

Introduction

One astonishing feature of the scientific enterprise is the role of a few extremely prolific researchers (Price, 1963). Thomson Reuters call them *Highly Cited Researchers* and they are listed and recognized per area. Based on another dataset, Scopus publications, Klavans & Boyack (2015) call them “superstars” and use them for large-scale studies of publication behaviour, thereby showing that superstars publishes less in isolated areas (retrieved using a clustering procedure), in dying areas, or in areas without an inherent dynamics. Highly productive and cited researchers tend to look for the new opportunities. Obviously, the highly productive researchers have to be taken into consideration for many reasons, both for science policy and for scholarly understanding of how the science system works.

Within bibliometrics there is a discussion on how to measure and to identify the superstars. Many current papers discuss the correlation between the various indicators developed for performance measurement. One of the stable outcomes is that there is a high correlation between the numbers of papers a researcher has published and the number of citations received (Bosquet & Combes, 2013). From that perspective, both indicators tend to measure the same attribute of researchers, as is actually materialized in the introduction of the H-index (Hirsch, 2005). Parallel, the discussion about impact has shifted from counting (field normalized) numbers of citations to more qualified types of citations and publications. As the progress of science rests on the huge amount of effort and publications, the number of real discoveries and path breaking new ideas is rather small. This has led to a different focus. Instead of counting publications and citations, the decisive difference is whether a researcher contributes to the small set of very highly cited papers. Different thresholds are deployed, from the top 1% or 10% of the highly cited papers or with the CCS method proposed by Schubert & Glänzel (1988). Only when reaching into these select set of papers that qualifies for citations above the x% level one can be considered as really having distinctive result that contributes to scientific progress. Increasingly, performance measures take this selectivity into account, and when calculating overall productivity and impact figures for researchers, papers (productivity) and citations (impact) are weighted differently depending on the impact percentile the paper belongs to (Sandström & Wold, 2015).

Of course, the question now comes up what a good publication strategy is – given this way of performance evaluation. Is publishing a lot the best way – or does that generally lead to normal

science, with low impact papers? The total number of citations received may still be large, but no top papers may have been produced. This is also the underlying idea of emerging movements in favour of ‘slow science’ like e.g., in the Netherlands; there the ‘science in transition’ movement (Dijstelbloem et al., 2014) was able to convince the minister of science and the big academic institutions to remove productivity as a criterion from the guidelines for the national research assessment (SEP). The underlying idea is that quality and not quantity should dominate – and that with all the emphasis on publications this has become corrupted.

However, others seem to see this differently. In his important work on scientific creativity, Simonton (2004) has extensively argued that (i) having a breakthrough idea is a low probability event that happens by chance, and therefore that (ii) the more often one tries, the higher the probability to have a ‘hit’ so now and then. There are also other contextual factors that may improve the chance for important results, but overall, the number of tries (publications) is the decisive variable. This also explains why Nobel laureates have so many more publications than normal researchers (Zuckerman, 1967; Sandström & Van den Besselaar, forthcoming). The more often you try (publish), the higher the probability that there is something very new and relevant, and atypical for the scientific community (Uzzi et al., 2013).

This brings us to the question whether there is a strong positive, or a negative relation between overall output (number of publications) and high impact papers. The answer of this question may inform our understanding of knowledge production and scientific creativity, but is also practically relevant for selection processes, and as explained above for research evaluation procedures: is high productivity a good thing, or a perverse effect and detrimental to the progress of science?

Methods and Data

In order to investigate this, we use the 74,000 WoS-publications 2008-2011 (with citations until 2014) of all researchers with a Swedish address using the following document types in databases SCI-E, SSCI and A&HCI: articles, letters, proceeding papers and reviews.

For identifying authors and keeping them separate we use a combination of automatic and manual *disambiguation* methods. An algorithm for disambiguating unique individuals was developed by Sandström & Sandström (2009), based on Soler (2007) and Gurney, Holdings & van den Besselaar (2012) and was found to proceed fast, although with minor manual cleaning methods. The deployed method takes into account surnames and first-name initials, the words that occur in article headings, and the journals, addresses, references and journal categories used by each researcher. There is also weighting for the normal publication frequency of the various fields.

As indicated, the data covers 74,000 articles and 195,000 author shares that have been judged to belong to Swedish universities or other Swedish organisations. In a few cases, articles from people who have worked both in Sweden and in one or more Nordic countries have been kept together, and articles have thus been included even if they came into being outside Sweden (the process of distinguishing names is thus carried out at Nordic level).

All articles by each researcher are ranked, based on received citations and according to the about 260 subject categories as specified in the Web of Science, and the articles are divided into CSS (Characteristic Scores and Scales) classes (0, 1, 2, 3). While measures based on percentile groups (e.g. top1% etc.) are arbitrarily constructed, CSS have some advantages concerning the identification of outstanding citation rates (Glänzel & Schubert, 1988). The CSS method is a procedure for truncating a sample (e.g., a subfield citation distribution) according to mean values from the low-end up to the high-end. Every group created using this procedure helps to identify papers that fulfil the requirements for being cited above the respective thresholds. In this paper we will use two levels, level CSS1 and CSS3, which in the

former case cover the 20%-25% most cited papers, and in the latter case the about 2%-3% of most cited papers: the “outstandingly cited papers” (Glänzel, 2011).

In this paper we will investigate the relation between quality and quantity in several different ways. We proceed in this way, as from a methodological perspective different options are open, without a convincing argument which one would be the better. By using a variety of methods, we avoid to produce results as artefacts of the method deployed.

(i) Firstly, we calculate the probability to have one, two or three and more top cited papers, given the productivity level. We calculate this for the health, i.e. medical sciences (about 15,000 researchers), where we classify these authors in several productivity classes. Class 1 has one publication in the four years period under study, class 2 has two, class 3 has three to four, class 4 has five to eight, class 5 has nine to sixteen, class 6 has seventeen to 31 publications, and finally class 7 covers researchers with 32 or more publications. Publications are integer counted, but citations are field normalized.

(ii) Secondly, we do a simple regression with the total number of (integer counted, IC) publications as the independent variable, and the (also integer counted) number of top cited publications in terms of one of the definitions as discussed above. Also, here citations are field normalized. We have here all researchers, without normalizing for field based productivity figures. As the total set of researchers is dominated by life and medical sciences and by natural sciences, and as these groups have comparable average publications and citations, we assume that this does not really influence the results. Under point four below, we introduce a way of taking field differences in productivity into account.

(iii) Thirdly, we do the same analysis as described above, but use fractional instead of full counting. This helps to investigate the effect of different ways of counting on the relations under study.

(iv) Fourthly, we move to the field-normalized (fractional counted) productivity, and calculate the relation between in this way defined productivity and having at least one publication in CSS1 respectively in CSS3. In the last analysis, we can provide an integrated analysis of all researchers across all fields, as we produced field normalized output counts. This is done with a method – Field Adjusted Production (FAP) based on Waring estimations – as initially developed by Glänzel and his colleagues (Braun, Glänzel & Schubert, 1990; Koski, Sandström & Sandström, 2011) during the 1980s. FAP is further explained and tested in Sandström & Sandström (2009). Basically, the method is used in order to compensate for differences between research areas concerning the normal rate of scholarly production. For this all journals in the Web of Science have been classified according to five categories (applied sciences, natural sciences, health sciences, economic & social sciences, and arts & humanities). Categorisation of journals into macro fields is based on Science Metrix classification of research into five major domains. Note that in some of the following analysis we will refrain from applying the Waring method, consequently, instead the analysis will be performed per scientific macro fields (for further information, see < <http://science-metrix.com/en/classification>>).

Results

(i) Does the probability of highly cited papers increase with productivity?

We calculated the number of top cited papers (CSS3) for each of the seven productivity classes. From this, Figure 1 was created. Clearly, the probability increases with productivity, and this is the case for 1, 2 and 3 or more papers in the CSS3 class. In fact, the relation is slightly different for the three criteria. The higher the criterion, the larger the effect is at the high end of the productivity distribution.

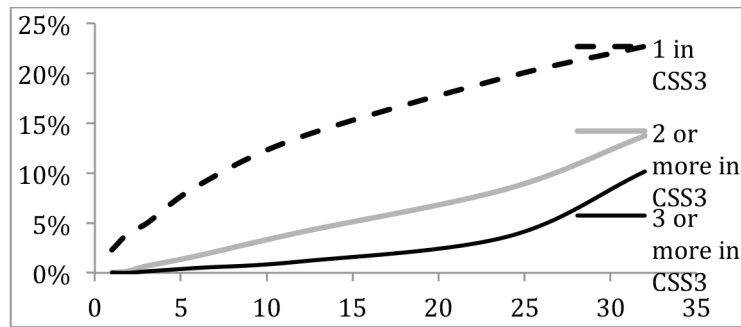


Figure 1. Share of papers in the CSS3 top cited class by productivity class.

(ii) *What is the effect of productivity on the number of highly cited papers?*

We have done a regression analysis with highly cited papers as dependent variable, and productivity as independent variable. We did the analysis for the various top cited classes. In the three figures below, we show the regression results. For papers in the top 1% of the cited papers (Figure 1) the correlation is about 0.5. For the CSS3, the top 10% of the cited papers, and the CSS1 classes, the correlations are 0.58, 0.78 and 0.88. The correlations are fairly high.

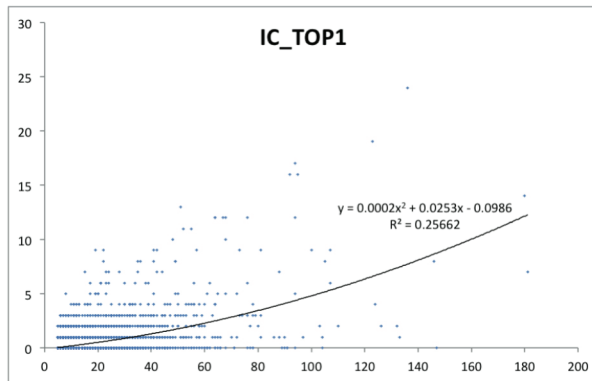


Figure 2. Top 1% of cited papers by total number of papers.

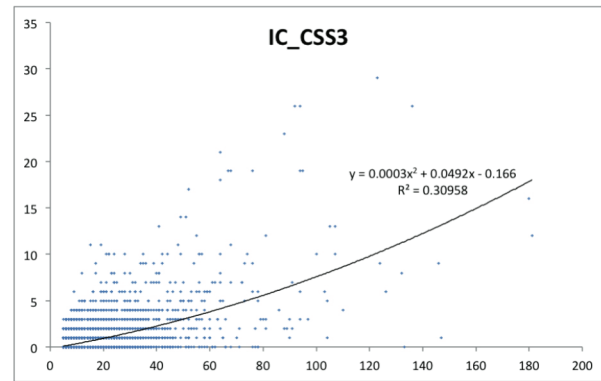


Figure 3. CSS3 cited papers by total number of papers.

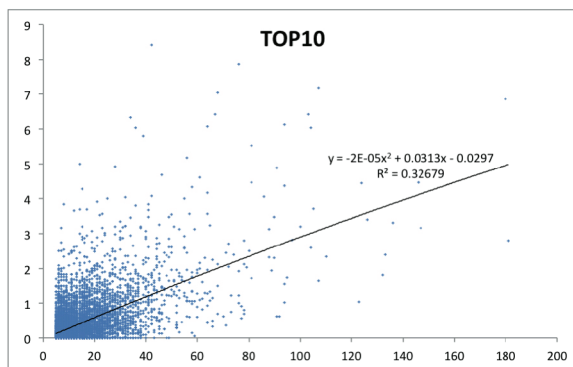


Figure 4. Top 10% of cited papers by total number of papers.

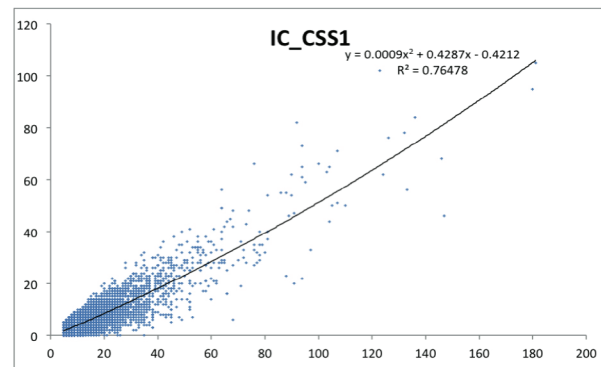


Figure 5. CSS1 cited papers by total number of papers.

Interestingly, the correlation becomes higher the lower the citation threshold. Why this is the case is not yet investigated. A possibility is that high productive researchers with top papers always have co-authors of these high cited papers who themselves are not highly productive. In that sense one also expects top cited authors in the lower productivity segments, reducing the explained variance. So probably, one should only include PIs in the analysis to avoid this effect. This could be the topic for a subsequent study.

One should realize that a small share of all authors produces most of the papers and of the highly cited papers. The 6.3% of most productive researchers (everybody above eleven publications in four years) are responsible for 37% of all papers and for 53% of the top 1% of the cited papers. Also this supports the idea that quantity makes a difference.

(iii) And the effect of fractional counted productivity on the number of highly cited papers?

We did the above analysis also using fractional counting of productivity. The patterns are the same, but the correlations are about .15 to .20 lower than in the full counted model. How this can be explained will be addressed in a coming paper. But also here, the 6.3% of the most productive authors are decisive: they have 46.8% of the fractional counted top 1% of the cited papers.

(iv) What is the effect of field adjusted production counting?

The relation between having at least one paper in CSS1 and total field normalized output is plotted in Figure 6, and as becomes obvious, the correlation is fairly high ($r = 0.79$), and not much smaller than in the above four where we did not use the field adjusted production (0.90, see Figure 5). The results here suggest that indeed the more papers someone publishes, the higher the probability of having a paper in the group of fairly good papers cited above the threshold of CSS1.

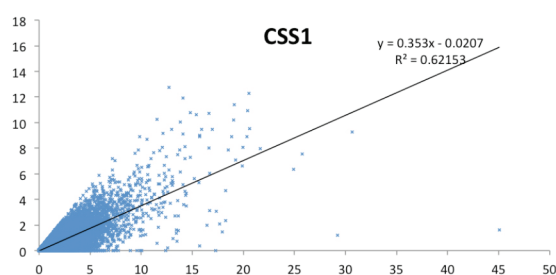


Figure 6. Fractionalized CSS1 by field adjusted production (all areas of science).

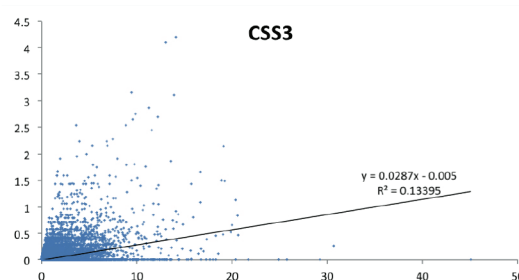


Figure 7. Fractionalized CSS3 by field adjusted production (all areas of science).

We also plot the relation between having at least one paper in the CCS3 (Figure 7), so in a much more narrow defined top, and field-normalized productivity, and although correlation is lower here, it is still considerable ($r = 0.37$). However, in the CSS3 case, the correlation when applying FAP is lower than the correlation without applying FAP (Figure 3), namely is 0.58. These differences need some further exploration.

The underlying distribution for the fields of Natural sciences and Medical and Life sciences are given in Table 1, which shows for seven distinct productivity categories the percentage of Swedish researchers in that category, the average number of papers published in a four-year period, the average fraction of paper production, and of course the percentage of researchers with at least one paper in CCS3.

As ‘field adjusted’ production (FAP) might be a rather abstract concept, we have translated it below for the various disciplines into ‘normal papers’. So, what is the relation between the number of papers produced (in a period of four years) and the probability of having a ‘top cited paper’ (in the top 2%-3% cited papers CSS3 class) during the period 2008-2014? This is a more sophisticated version of the analysis presented in section (i) above. As we clearly see in Table 2, the higher the number of papers, the more likely that one has a paper that ends up to be an outstandingly cited paper. Actually, the increase is rather steep and one may say that in most disciplines only with some ten papers in the period under consideration, there is a good chance of having a top paper. The humanities have a different pattern, as with a production of five papers one has the highest chance of reaching the top.

Table 1. CSS3 papers by production levels, Health sciences and Natural sciences

Category	Medical and life sciences				Natural sciences			
	researchers	Mean P	Frac P	CSS3	researchers	Mean P	Frac P	CSS3
1 (1 paper)	40.8%	1	0.2	0.03	9,0%	1	0.2	0.02
2 (2 papers)	16.92%	2	0.4	0.06	16,3%	2	0.5	0.05
3 (>2-4)	17.08%	3.4	0.7	0.10	17,4%	3.4	0.9	0.10
4 (>4-8)	13.36%	6.1	1.3	0.21	13,7%	6.1	1.6	0.21
5 (>8-16)	7.23%	11.6	2.4	0.44	8,3%	11.5	2.8	0.40
6 (>16-32)	3.36%	22.3	4.4	1.05	4,1%	22.0	4.7	0.87
7 (>32)	1.18%	50.5	8.8	3.45	1,2%	47.6	9.8	2.68
Average		4.3	0.9	0.17		4.6	1.1	0.17

Data for this table is built on publications from 37,114 researchers.

Table 2: Probability of one outstanding paper (CSS3) at different levels of production.

Average # of publications	Discipline					
	Class	Natural	Health	Applied	Ec & Soc	Hum
1	1	5%	7%	7%	6%	9%
2	2	11%	13%	13%	13%	8%
3	3	20%	21%	21%	24%	25%
6	4	31%	34%	33%	34%	33%
11	5	49%	54%	53%	55%	33%
20	6 / 7	61%	80%	66%	83%	
38	7			88%		
46	7	83%				
49	7		93%			

Note: Data for this table consist of $\approx 190,000$ article shares with <40 authors per paper. The numbers of publication are the field-specific averages per productivity class (for more information, see Table 1).

Conclusions

As the above results show, there is not only a strong correlation between productivity (number of papers) and impact (number of citations), that also holds for the production of high impact papers: the more papers, the more high impact papers. In that sense, increased productivity of the research system is not a perverse effect of output oriented evaluation systems, but a positive development, as it strongly increases the occurrence of breakthroughs and important inventions (c.f. Uzzi et al., 2013). The currently upcoming discussion that we are confusing quality with quantity therefore lacks empirical support. As we deployed a series of methods, with results all pointing in the same direction, the findings are not an artefact of the selected method.

The analysis also gives an indication of the output levels that one may strive at when selecting researchers for grants or jobs.

We also plan some future work: Firstly, we plan to extend the analysis to some other countries, which of course requires large-scale disambiguation of author names. Secondly, we will in a next version control for number of co-authors, and for gender. The former relates to the discussion about team size and excellence, the latter to the ongoing debate on gender bias and gendered differences in productivity. Thirdly, the aim is to concentrate on principle investigators, and remove the incidental co-authors with low numbers of publications, as they may seem to be high impact authors at the lower side of the performance distribution. This all should lead to a better insight in the relation between productivity and impact in the science system.

References

- Bosquet, C. & Combes, P-P. (2013). Are academics who publish more also more cited? Individual determinants of publication and citation records. *Scientometrics*, 97: 831-857.
- Braun, T., Glänzel, W. & Schubert, A. (1990). Publication productivity: from frequency distributions to scientometric indicators. *Journal of Information Science*, 16: 37-44.
- Dijstelbloem, H., Huisman, F., Miedema, F. & Mijndhardt, W. (2014). Science in Transition Status Report: Debate, Progress and Recommendations. <http://www.scienceintransition.nl/wp-content/uploads/2014/07/Science-in-Transition-Status-Report-June-2014.pdf>.
- Glänzel, W. (2011). The application of characteristic scores and scales to the evaluation and ranking of scientific journal. *Journal of Information Science*, 37(1): 40-48.
- Glänzel, W. & Schubert, A. (1988). Characteristic scores and scales in assessing citation impact. *Journal of Information Science*, 14: 123-127.
- Gurney, T., Horlings, E. & van den Besselaar, P. Author disambiguation using multi-aspect similarity indicators. *Scientometrics*, 91: 435-449.
- Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *PNAS*, 102(46): 16569-16572.
- Klavans, R. & Boyack, R.W. (2015). Scientific superstars and their effect on the evolution of science. http://www.enid-europe.org/conference/abstract%20pdf/Klavans_Boyack_superstars.pdf.
- Koski, T., Sandström, E. & Sandström, U. (2011). Estimating research productivity from a zero-truncated distribution. Paper to the 2011 ISSI Conference in Durban.
- Price, D.J.S. (1963). *Little Science, Big Science*. New York: Columbia University Press.
- Sandström, U. & Sandström, E. (2009). The field factor: towards a metric for academic institutions. *Research Evaluation*, 18(3): 243-250.
- Sandström, U. & van den Besselaar, P. (2015). *Before the prize: Nobel Prize laureates recognition by their scientific community*. Manuscript in preparation.
- Sandström, U. & Wold, A. (2015). Centres of Excellence: reward for gender or top-level research? In B. Bjorkman & B. Fjaestad (Eds.). *Thinking Ahead: Research, Funding and the Future* (pp. 69-91). Stockholm, Makadam Publ.
- Simonton, D.K. (2004). *Creativity in Science: Chance, Logic, Genius, and Zeitgeist*. New York: Cambridge Univ Press. [Reprinted 2008].
- Soler, J-M. (2007). Separating the articles of authors with the same name. *Scientometrics* 72 (2): 281-290. DOI: 10.1007/s11192-007-1730-z.
- Uzzi, B., Mukherjee, S., Stringer, M. & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, 342, 468-472.
- Zuckerman, H. (1967). Nobel laureates in science: Patterns of productivity, collaboration, and authorship. *American Sociological Review*, 32 (3): 391-403.

On Decreasing Returns to Scale in Research Funding

Philippe Mongeon¹, Christine Brodeur¹, Catherine Beaudry² and Vincent Larivière³

¹ *philippe.mongeon@umontreal.ca; christine.brodeur@umontreal.ca;*

Université de Montréal, École de bibliothéconomie et des sciences de l'information,

C.P. 6128, Succ. Centre-Ville, Montréal, QC. H3C 3J7 Canada

² *catherine.beaudry@polymtl.ca*

École Polytechnique de Montréal, Département de mathématiques et de génie industriel,

C.P. 6079, Succ. Centre-Ville, Montréal, QC. H3C 3A7 Canada

Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), CP 8888, Succ. Centre-Ville,
H3C 3P8 Montreal, Qc. (Canada)

³ *vincent.lariviere@umontreal.ca*

Université de Montréal, École de bibliothéconomie et des sciences de l'information, C.P. 6128, Succ.

Centre-Ville, H3C 3J7 Montreal, Qc. (Canada) and

Université du Québec à Montréal, Centre Interuniversitaire de Recherche sur la Science et la Technologie
(CIRST), Observatoire des Sciences et des Technologies (OST), CP 8888, Succ. Centre-Ville, H3C 3P8
Montreal, Qc. (Canada)

Abstract

In most countries, basic research is supported through governmental research councils that select, after peer review, the individuals or teams that will receive funding. Unfortunately, the number of grants these research councils can allocate is not infinite, and many researchers (45% in Quebec) are not able to obtain any funding. A small minority of those who do get funded account for the majority of the available funds. However, it is unknown whether or not this is an optimal way of distributing available funds. The purpose of this study is to measure the relation between the amount of funds given to 14,103 individual Quebec's researchers over a fifteen year period (1998-2012) and the total outcome of their research in terms of output and impact from 2000 to 2012. Our results show that both in terms of the quantity of papers produced and of their scientific impact, the concentration of research funding in the hands of a so-called 'elite' of researchers generally produces diminishing returns.

Conference Topic

Science policy and research assessment

Introduction

In most countries, basic research is supported through governmental research councils that select, after peer review, the individuals or teams that will receive funding. Unfortunately, the number of grants these research councils can allocate is not infinite. For example 20% to 45% of Quebec's researchers, depending on the discipline, had no external funding between 1999 and 2006 (Larivière et al., 2010). National scientific agencies, including the National Science Foundation (NSF – United States) and Natural Science and Engineering Research Council (NSERC – Canada), also tend to give fewer grants of a higher value, which leads to high rejection rates (Joós, 2012; NSERC, 2012; NSF, 2013). In Canada, 10% of the researchers funded by the Social Sciences and Humanities Research Council (SSHRC) accumulate 80% of available funds, 10% of those funded by the Canadian Institutes of Health Research (CIHR) obtain 50% of the funds, and 10% of those funded by the NSERC accumulate 57% of the funds.¹ The situation is similar in Quebec where we combine funding from the national

¹ Data compiled by the Observatoire des Sciences et Technologies (OST) using results of competition for each of the councils, and the *Almanac of Post-Secondary Education in Canada, of the Canadian Association of University Teachers*.

and provincial agencies: 20% of the researchers getting 80% of the funds in social sciences and humanities (SSH), 50% of the funds in health, and 57% of the funds in natural sciences and engineering (NSE) (Larivière et al., 2010). With a few researchers receiving most of the funds available and many not receiving any, it seems legitimate to ask whether this concentration of funds leads to better collective gains than funding policies that promote a more even distribution of funding. The aim of this study is to provide a partial answer to this question, by linking the amount of funding obtained by Quebec's scientists with their research productivity and impact.

Even though the funding of science theoretically plays a substantial role in scientific discoveries, its relation to outcomes has not been extensively researched. McAllister and Wagner (1981) observed a linear relationship between funding and output at the institution level. A few years later, Moed et al. (1998) found that departments of Flemish universities with the most funding actually had a decrease in publications. Other studies (e.g., Heale et al., 2004 and Nag et al., 2013) investigated the relation between the amount of funding and the research output of individual researchers. They reported that one of the strongest determinants of the number of publications was the amount of funding, although an increase in funds did not yield a proportional increase in the number of articles. Thus, there are decreasing returns to scale. Others have found that productivity is only weakly related to funding (Fortin & Currie, 2013), and that publications do not increase linearly with the amount of funding but rather appears to reach a plateau (Berg 2010). On the whole, while most studies—unsurprisingly—found a positive relationship between inputs and outputs, very few have looked at decreasing returns to scale associated with the concentration of research funding. Nicholson and Ioannidis (2012) found that only a minority (about 40%) of all researchers eligible to NIH funding who published highly cited articles (1000 citations or more) actually received such funding. Previous studies found that funded researchers publish more (Gulbrandsen & Smeby, 2005) and are more cited (Zhao, 2010; Jowkar, 2011; Campbell et al., 2010; van Leeuwen et al., 2012) than those who do not receive any funding.

This study aims to contribute to this debate, by analyzing the research output and impact of all of Quebec's researchers from all disciplines over a period of 15 year. More specifically, it aims at answering two questions: 1) how does the research productivity and scientific impact of individual researchers vary with the amount of funding they receive? 2) Is this variation similar in the three general fields of science that are health, natural sciences and engineering, and social science and humanities?

Methods

Data on funding for all Quebec's academic researchers from 1998 to 2012 were obtained from the Information System on University Research, an administrative database from the Quebec provincial government that covers all funded research in Quebec's universities. Researchers were divided in three broad research disciplines: Social Sciences and Humanities (SSH), Natural Sciences and Engineering (NSE) and Health according to the discipline of their university department. Some were put in two different disciplines (N=169), and those for whom the discipline was not known and not found were excluded (N=263). The number of researchers in each field is shown in table 1. For each researcher, we calculated the total amount of funding received from the three main funding agencies in Quebec (FRQSC [SSH], FRQNT [NSE] and FRQS [health]) and Canada (SSHRC [SSH], NSERC [NSE] and CIHR [Health]). The total funds attributed for each projects were divided equally by the number of researchers on the application, each of them receiving an equal share. Other sources of funding were not taken into account. Publication data for each researcher from 2000 to 2012 were obtained from Thomson Reuters' Web of Science. Since citations take time to accumulate, they were counted up to the end of 2013.

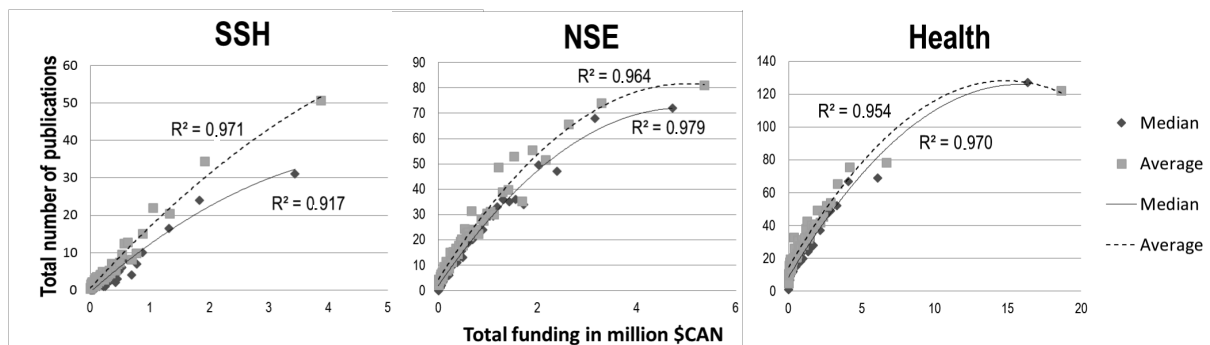
Table 1. Number of Quebec's researchers by field

<i>Field</i>	<i>Number of researchers</i>	<i>Funded</i>		<i>Not funded</i>	
		<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
SSH	6,229	3,869	62.1%	2,360	37.9%
NSE	3,244	2,647	81.6%	597	18.4%
Health	4,630	2,666	57.6%	1,964	42.4%
Total	14,103	9,182	65.1%	4,921	34.9%

Similarly to Berg (2010), we divided researchers in bins of equal size (50 researchers per bin), except for the bin regrouping researchers who did not receive any funding (see table 1 for the number of researcher in each field who did not receive funding). For each bin, we calculated the average and median amount of funding received. Then we calculated the average and median of four indicators used to measure the research outcome: the total count of articles, the fractional number of articles, the total number of citations and the average relative citations (ARC).

Results

Figure 1 and Figure 2 provide the mean and median number of papers of researchers, using both full (Figure 1) and fractional counting (Figure 2), as a function of total funding received. For each bin for each discipline and each indicator, the average is higher than the median, implying a skewed distribution of the data. The high values of R^2 in both figures indicate that the number of publications is strongly linked to the amount of funding received by researchers. The best fit line for each domain is a quadratic equation which suggests diminishing returns. For example, the median number of publications of researchers in NSE who received about \$5 million is about 72 (and 19 for fractional count), while those who receive \$2.5 million published a median number of 47 articles (13 for fractional count). Thus, doubling the funding does not seem to double the output. In Health, the most funded bin received almost three times more funding than the second most funded one, but published only two times more articles. Furthermore, in health, the apex is reached within the data range, which shows that a decline in production could be associated with higher levels of funding. On the whole, the correlation between funding and publications appears to be strong in all fields with values of R^2 higher than 0.91, but for each domain and calculation method, a rapid growth in the number of publications is observed for smaller amounts received and is followed by a slower growth as funding increases. However, this effect is less apparent for the total number of publications in SSH.

**Figure 1. Full number of publications as a function of the amount of funding received.**

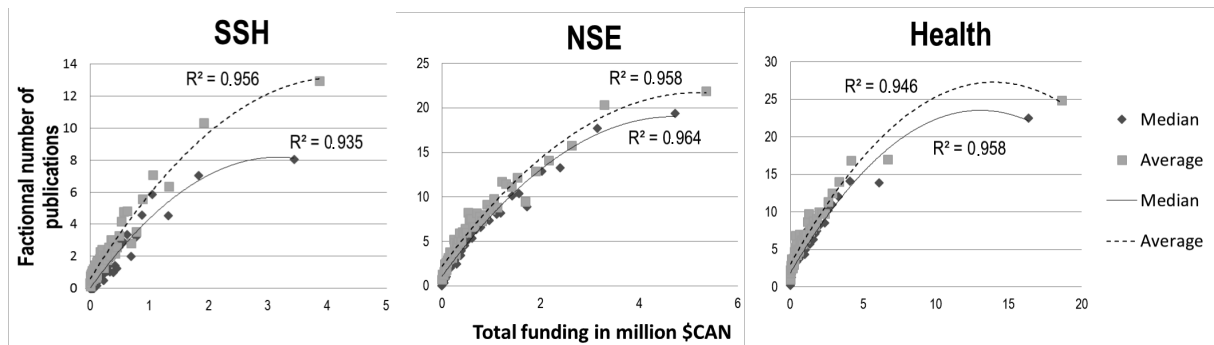


Figure 2. Fractional number of publications as a function of the amount of funding received.

Figure 3 shows the relationship between raw citations and funding received; the best-fit line is also a quadratic equation suggesting decreasing returns to scale in scientific impact. Similar to publications, the relation between the average of relative citations and the amount of funding (Figure 4) is weaker than for the previous indicators, with R^2 between 0.4 and 0.9. The nature of the relation is also different, the best-fit line being a power function, except for the median in SSH and the average in NSE, which are quadratic function. The power function indicates decreasing returns: the average relative of citations keeps increasing when increasing the total of funding, but not proportionally. For both impact indicators, we observe a trend similar to that observed for the number of publications. While the impact of papers published increase rapidly for funding of less than approximately \$2 million in NSE and \$5 million in health, the total number citations increase at a much slower pace once this threshold is met. Here, SSH are the exception, with the total number of citation seemingly increasing more rapidly for highly funded researchers. For field-normalized citations, the impact remains almost the same for all fields after a threshold of approximately \$1 million is met.

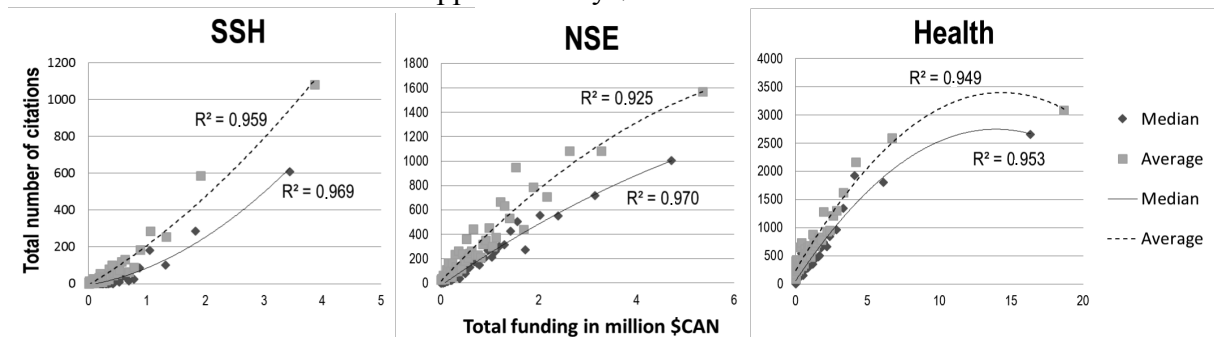


Figure 3. Total number of citations as a function of the amount of funding received.

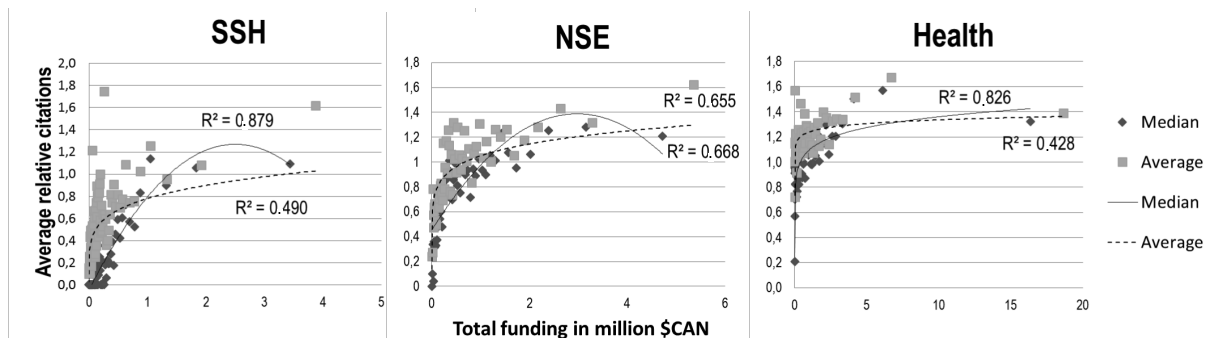


Figure 4. Average relative citations as a function of the amount of funding received

Discussion and conclusion

Based on our observations, funding is strongly linked to productivity and impact of individual researchers, but there are decreasing returns to scale for all of the indicators measured, except for the total citation count in SSH. This suggests that, even though more funding does in general lead to a higher number of publications, giving bigger grants to fewer individuals may not be optimal. If maximum output is the objective, then giving smaller grants to more researchers seems to be a better policy. In terms of scientific impact, the quickly reached plateau indicates that increasing funding has a very small impact on relative citations. Again, if the goal of research funding is to generate research that has a greater impact, giving grants to more researchers seems to be a better decision.

According to our results, SSH seem to be an exception, showing very little decreasing returns to scale. However, this could be explained by the fact that some research specialties in SSH (e.g., psychology and geography) have publication practices that are similar to those in NSE or Health. A closer look at the data shows that some researchers in psychology and geography tend to be both more funded – since they are often funded by the health and natural sciences funding agencies respectively – and more prolific than those in other field. Twenty-three (23) of the 50 most funded researchers and 33 of the 50 most prolific researchers are in those two fields, while they were 10 out of 50 in a randomly selected bin of researchers with less funding. Thus, the lower decrease in return of research funding in SSH could potentially be explained by an overrepresentation psychology and geography researchers in the highly funded bins, and their underrepresentation in less funded ones.

One of the many potential explanations for these decreasing returns is the high cost of equipment and infrastructures. Some research projects may simply not be possible without these initial investments, which do not necessarily lead to more output. Furthermore, while receiving funding does provide researchers with the means to carry on their research projects, it does not guarantee that they will succeed at achieving publishable results. Research grants are sometimes used as a performance indicator, which encourages researchers to apply for more grants (Hornbostel, 2001) that they might not necessarily need. This could lead to an inefficient use of the funds received (Sousa, 2008). Another explanation could be that researchers receiving larger grants may not participate directly on all the work funded with those grants (Boyack & Jordan 2011)

Some limitations of this study should be acknowledged. We did not control for other factors that can have an impact on a researcher's productivity (e.g., team size, academic age or gender), so further research may want to take into account such factors, as well as sources of funding other than government grants. Also, some of the potential outcome of funding and research cannot be measured with bibliometric indicators (e.g., the number of students trained and social outcomes). The funding received is sometimes linked to a particular project, and further research could aim at comparing outcomes of funded projects specifically. Another limit might be the lower coverage of SSH publications in the Web of Science, since researchers in SSH tend to publish in local journals or to publish books. Finally, as discussed above, the division of researchers in three broad disciplines might be problematic, especially for SSH. A more precise clustering of researchers based on research topic could provide better results and a clearer understanding of the phenomenon of decreasing returns of research funding.

In sum, both in terms of the quantity of papers produced and of their scientific impact, the concentration of research funding in the hands of a so-called 'elite' of researchers generally produces diminishing returns. In a context where financial resources devoted to research are declining in constant dollars, it is important to ask whether the way funding is allocated is optimal. Our numbers show that it is not the case: a more egalitarian distribution of funds would yield greater collective gains. It should be understood that the main determinant of

scientific production is not so much the money invested, but, rather the number of researchers' at work and, by funding a greater number of researchers, we increase the overall research productivity. Research policies that concentrate financial resources also seem to forget that there is a certain degree of serendipity associated with scientific discoveries, and by funding the work of many researchers as possible, we increase the likelihood that some of them make major discoveries.

Acknowledgments

This work was funded by the Canada Research Chairs program, and by the Social Sciences and Humanities Research Council of Canada.

References

- Berg, J. (2010). Measuring the scientific output and impact of NIGMS grants. *NIGMS Feedback Loop Blog*. Retrieved June 15, 2015 from: <http://loop.nigms.nih.gov/2010/09/measuring-the-scientific-output-and-impact-of-nigms-grants/>
- Boyack, K. W. & Jordan, P. (2011). Metrics Associated with NIH Funding: A High-Level View. *Journal of the American Medical Informatics Association*, 18(4), 423–431.
- Campbell, D., Picard-Aitken, M., Côté, G., Caruso, J., Valentim, R., Edmonds, S., & Archambault, E. (2010). Bibliometrics as a performance measurement tool for research evaluation: the case of research funded by the National Cancer Institute of Canada. *American Journal of Evaluation*, 31(1), 66-83.
- Fortin, J. M. & Currie, D. J. (2013). Big science vs. little science: how scientific impact scales with funding. *PLoS ONE*, 8(6), e65263.
- Gulbrandsen, M. & Smeby, J.C. (2005). Industry funding and university professors' research performance. *Research Policy*, 34(6), 932-950.
- Heale, J.P., Shapiro, D. & Egri, C.P. (2004). The determinants of research output in academic biomedical laboratories. *International Journal of Biotechnology*, 6(2-3), 134-154.
- Joós, B. (2012). NSERC's discovery grant program: disquieting changes & why they matter to Canadian science. *CAUT Bulletin*, 59(1).
- Jowkar, A., Didegah, F. & Gazni, A. (2011). The effect of funding on academic research impact: a case study of Iranian publications. *Aslib Proceedings*, 63(6), 593-602.
- Larivière, V., Macaluso, B., Archambault, É. & Gingras, Y. (2010). Which scientific elites? On the concentration of research funds, publications and citations. *Research Evaluation*, 19(1), 45-53.
- McAllister, P. R. & Wagner, D. A. (1981). Relationship between r-and-d expenditures and publication output for United-States colleges and universities. *Research in Higher Education*, 15(1), 3-30.
- Moed, H. F., Luwel, M., Houben, J. A., Spruyt, E. & Van Den Berghe, H. (1998). The effects of changes in the funding structure of the Flemish universities on their research capacity, productivity and impact during the 1980's and early 1990's. *Scientometrics*, 43(2), 231-255.
- Nag, S., Yang, H., Buccola, S. & Ervin, D. (2013). Productivity and financial support in academic bioscience. *Applied Economics*, 45(19), 2817-2826.
- Nicholson, J. M. & Ioannidis, J. P. A. (2012). Research grants: Conform and be funded. *Nature*, 492(7427), 34-36.
- NSERC. (2012). *2012 Competition Statistics Discovery Grants Program*. Retrieved June 15, 2015 from: http://www.nserc-crsng.gc.ca/_doc/Funding-Financement/DGStat2012-SDStat2012_eng.pdf.
- NSF. (2013). *Summary Proposal and Award Information (Funding Rate) by State and Organization*. Retrieved June 15, 2015 from: <http://dellweb.bfa.nsf.gov/awdfr3/default.asp>.
- Sousa, R. (2008). Research funding: less should be more. *Science*, 322(5906), 1324-1325.
- van Leeuwen, T.N. & Moed, H.F. (2012). Funding decisions, peer review, and scientific excellence in physical sciences, chemistry, and geosciences. *Research Evaluation*, 21(3), 189-198.
- Zhao, D. Z. (2010). Characteristics and impact of grant-funded research: a case study of the library and information science field. *Scientometrics*, 84(2), 293-306.

How Many is too Many?

On the Relationship between Output and Impact in Research

Vincent Larivière¹ and Rodrigo Costas²

¹ *vincent.lariviere@umontreal.ca*

Université de Montréal, École de bibliothéconomie et des sciences de l'information, C.P. 6128, Succ. Centre-Ville, H3C 3J7 Montréal, Qc. (Canada) and Université du Québec à Montréal, Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Observatoire des Sciences et des Technologies (OST), C.P. 8888, Succ. Centre-Ville, H3C 3P8 Montreal, Qc. (Canada)

² *rcostas@cwts.leidenuniv.nl*

Leiden University, Center for Science and Technology Studies, Wassenaarseweg 62A, 2333 AL Leiden (The Netherlands)

Abstract

Over the last few decades, the massification of quantitative evaluations of science and their institutionalisation in several countries has led many researchers to aim at publishing as much as possible. This paper assesses the potential adverse effects of this behaviour by analysing the relationship between individual researchers' productivity and their proportion of highly cited papers. In other words, does the share of an author's top 1% most cited papers increase, decrease or remain stable, as her number of total papers increase? Using a large dataset of disambiguated researchers (N= 25,994,021) over the 1980-2012 period, this paper shows that the higher the number of papers a researcher publishes, the more likely they are amongst the most cited in their domain. This relationship was stronger for older cohorts of researchers, while decreasing returns to scale were observed in some domains for more recent cohorts. On the whole, these results suggest that at the macro-level, the culture of publishing as many papers as possible did not yield to adverse effects in terms of impact, especially for older researchers. For such researchers, who have had a long period of time to accumulate scientific capital, there can never be too many papers.

Conference Topic

Science Policy and Research Assessment

Introduction

In the second half of the 20th Century, but even more so over the last few decades, evaluations have become widespread in various spheres of society (Dalher-Larsen, 2011). Although scientific research has long been exempt from external evaluations thanks to Vannevar Bush and post WWII non-interventionist science policy, it has always been assessed internally through peer review. These means of evaluating research and researchers have, however, slowly changed since the 1980s, when researchers and administrators became aware of the roles that bibliometric analyses could play in such evaluations. Quantitative publication and citation analyses gained even more importance in the 2000s (Cameron, 2005), when tools for assessing individual researchers' output and impact became widespread. While in some cases, these methods have been developed to complement peer review in the allocation of research funding—such as the BOF-key in Flanders (Belgium) (Debackere & Glänzel, 2004), the Research Assessment Exercise/Framework in the UK—in other settings, these quantitative evaluations of research have become the main mean through which research is assessed and funded (Sörlin, 2007). Various publication-based and citation-based funding models can be found in Australia, Norway, Denmark, Sweden and Finland—and translates as the currency through which academic exchanges of tenure, promotion and salary raises are made (e.g. Fuyono & Cyranoski, 2006).

While there has always been subliminal bibliometrics performed through peer evaluation—as reviewers were skimming through reviewees' CVs through the process—the massification of

evaluations and their institutionalisation led many researchers and institutions to put large emphasis on the number of papers they published. This has led to adverse effects (Binswanger, 2015; Frey & Osterloh, 2006; Haustein and Larivière, 2014; Weingart, 2005). Indeed, like any social group, researchers are prone to change their behaviour once the rules of the games become explicit or what is expected from them; phenomenon that could be referred to as the Hawthorne effect (Gillespie, 1993), or to Goodhart (1975) or Campbell's laws (1979). As most evaluations and rankings are first based on numbers of published papers, this has created incentives for researchers to *author as many papers as possible*. In Australia (Butler, 2004), where publications counts were used without differentiating between publication venue or citations received, researchers have been found to increase their numbers of publications in journals with high acceptance rates and lower impact. Along these lines, the h-index, which together with the Impact Factor, is likely the most popular bibliometric indicator in the scientific community, is largely determined by numbers of papers published than on citations (Waltman & van Eck, 2012).

Within this context, researchers have adopted many publication strategies. While some researchers focus on publishing few, high-quality papers—e.g. 'selective' (Costas & Bordons, 2007) or 'perfectionists' (Cole & Cole, 1973)—others publish as many papers as possible, without not all of them necessarily being of high quality—e.g. 'prolific scientists' (Cole & Cole, 1973) or 'big producers' (Costas & Bordons, 2008)). However, little is known on the publication strategy that yields the highest results in terms of impact. In order to better understand the relationship between productivity and impact, this paper compares, for a large dataset of disambiguated researchers (N= 25,994,021), their total number of papers with the proportion of these papers that made it to the top 1% most cited of their field. Thus, this paper aims at answering the following key question: Does an authors' share of top papers start to decrease with a certain number of papers published? Or is it stable, as production and impact are two distinct dimensions of scientific activity. In other words, *how many is too many?* What is the probability for an author to publish top cited papers relate to the number of papers published? A good analogy for this is archery: if an archer throws one arrow, what is the probability that it hits the center of the target? Does an increase in the number of arrows thrown leads to an increase in the proportion of arrows hitting the center of the target?

Two opposite hypotheses could be made. The first one would be that authors with just 'average' production—rather than low or high production— are the ones more likely to publish top cited papers, as these authors, perhaps, focus more on the 'quality' of their output than just on quantity (i.e. selective scholars). The second hypothesis would be that, it is the authors with very high number of papers who, on average, publish the highest proportion of top cited papers. This hypothesis would be on agreement with the theory of Merton's cumulative advantages (Merton, 1968), and supported by empirical work in the sociology of science (Cole & Cole, 1973). Similarly, in a Bourdieusian framework, the main goal of a researcher is to increase its rank in the scientific hierarchy and gain more scientific capital (Bourdieu, 2004). If publishing a high number of scientific papers and being abundantly cited are the ways through which researchers can reach this goal, then they will adapt their behaviour to reach these evaluation criteria.

This focus on publishing as many papers as possible—often referred to as 'salami slicing'—has been long discussed (e.g. Abraham, 2000; Jefferson, 1998). However, only a few authors have analysed the effect of 'salami slicing' on papers' citations. For instance, Bornmann and Daniel (2007) have shown, for a small sample of PhD research projects in biomedicine (N=96), that an increase in the number of papers associated with a project lead to an increase in the total citation counts of papers associated with the projects. However, they do not show whether the impact of each paper taken individually increases with the number of papers published. Similar to this study, Hanssen and Jørgensen (2015) analysed the effect of

‘experience’ on papers’ citations; experience being defined as the author’s previous number of publications. Drawing a sample of papers in transportation research (N=779) they show that experience is a statistically significant determinant of individual papers’ citations, although this increase becomes marginal once a certain threshold is met in terms of previous papers published.

Methods

This paper uses Thomson Reuters’ Web of Science (WoS) for the period 1980-2012. Only journal articles are included. Given that the units analysed in this paper are individual researchers, we used the disambiguation algorithm developed by Caron & van Eck (2014) to identify the papers of individual researchers. On the whole, the algorithm managed to attribute papers to 25,994,021 individuals, which were divided into seven cohorts based on the year of their first publication (Table 1).

Table 1. Number of disambiguated researchers per cohort

Year of first publication	Number of researchers
<=1985	3,574,667
1986-1990	2,733,002
1991-1995	3,282,421
1996-2000	3,810,652
2001-2005	4,310,886
2006-2011	6,930,289
>=2012	1,352,104

As we want to assess researchers’ contribution to research that has the highest impact, we isolated for each discipline the top 1% most cited papers published each year (normalized by WoS subject categories). Citations are counted until the end of 2013, and exclude self-citations. The broad disciplines used are those of the 2013 Leiden ranking which are based on the assignment of WoS Subject Categories to five main domains (CWTS, 2013). Figures in the paper presents classes of numbers of papers in which there are at least 100 researchers.

Results

Figure 1 presents, for the oldest cohort studied—researchers who have published their first paper before 1985—the relationship between the number of papers throughout their career and the proportion of those papers that made it to the top 1% most cited. For any specific number of papers, the expected value of top 1% papers is, as one might expect, 1%. Researchers for all five domains have one thing in common: authors with very few papers are, on average, much less likely to publish high shares of top 1% most cited papers. For *Biomedical and health sciences* and for *Social sciences and humanities* we observe a continuous increase in authors’ proportion of top papers as their overall number of papers increases. For *Life and earth sciences* the share of papers does increase with the number of papers, until about 10 papers where they starts to oscillate, although in general an increasing pattern is still observed, especially after 40 papers. Perhaps the most deviant pattern is found in *Mathematics and computer science* where for just for the very low levels of production there is an increase in the share of highly cited publications, but this share decreases between

4 and 20 papers. It then starts to increase again for higher numbers of papers, despite important fluctuations. Natural sciences and engineering follow a similar pattern, with a decrease in the share of top papers between 6 and 30 papers, followed, in this case, by a clear increase until very high levels of productivity.

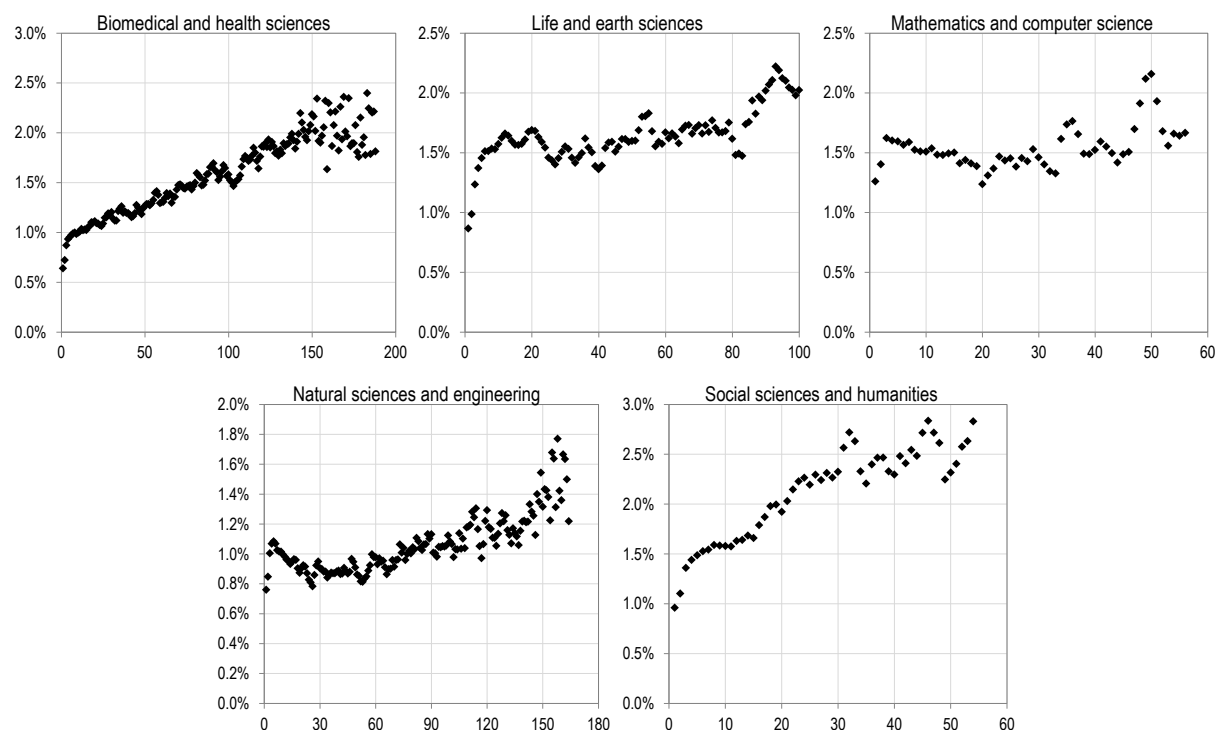


Figure 1. Proportion of top 1% most cited papers (y axis), as a function of the number of papers published (x axis), for the cohort of researchers who have published their first paper before 1985, by domain. Only classes of numbers of papers with 100 researchers or more are shown.

When researchers who have published their first paper between 2006 and 2011 are considered, different pattern are observed (Figure 2). For *Biomedical and health sciences* there is an increase in the share of highly cited publications up to around 15 publications, when some important fluctuations—or certain decreasing returns to scale—start to appear. A similar pattern is observed for the *Life and earth sciences* with the variability starting from levels of production of around 10 publications although, in this case, a decrease is clearly observed. For the other domains the pattern tends to be clearly increasing, although oscillations are also observed for the higher levels of production, which could also be seen as decreasing returns to scale. For the other three domains, there is clearly an increase in the share of top papers as the number of papers increases. However, we also observe for these three fields a decrease at very high levels of productivity.

An important characteristic of this cohort is that it got socialized to research recently—when the evaluation culture was more present—which might explain why they might be more prone to try to publish as much as possible. However, the drop in the share of top papers observed in each domain—although at different levels of productivity—suggests that these academically-younger scholars struggle to keep impact high once a certain threshold is met. This might be due to the fact that these scholars have not yet secured permanent or tenure positions and, thus, might feel that they cannot be as selective as older scholars who might choose their collaborators more easily.

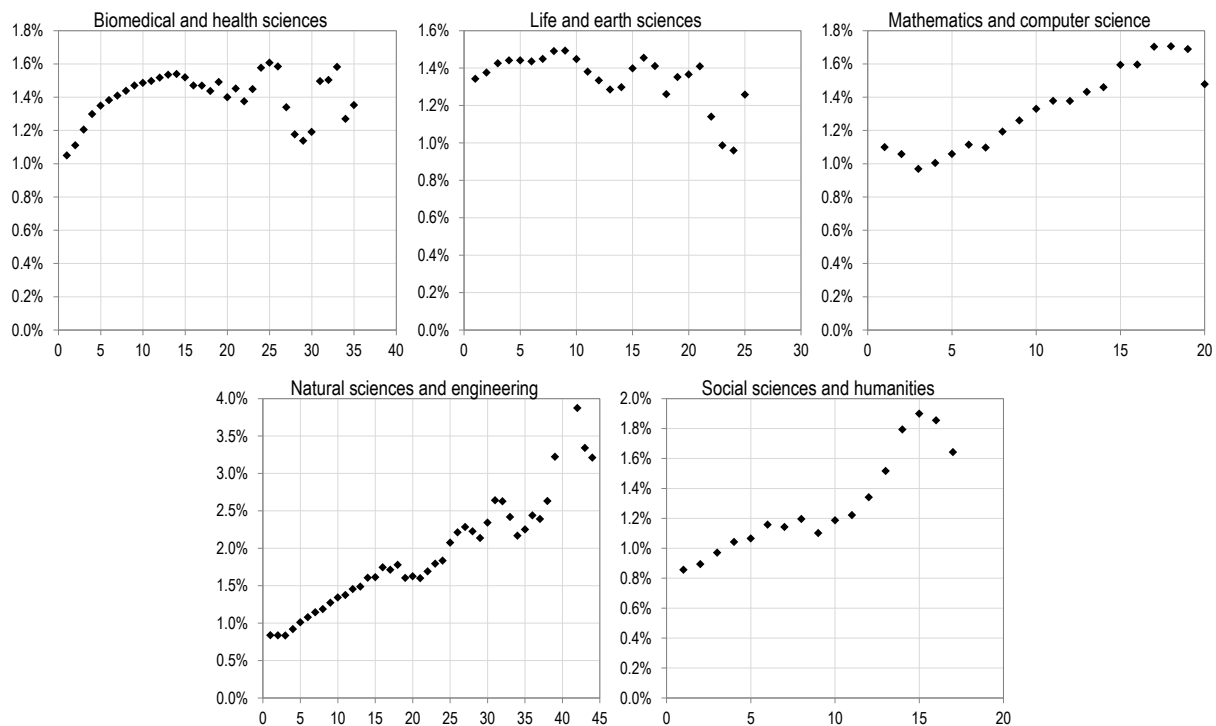


Figure 2. Proportion of top 1% most cited papers, as a function of the number of papers published, for the cohort of researchers who have published their first paper between 2006 and 2011, by domain. Only classes of numbers of papers with 100 researchers or more are shown.

Discussion and Conclusion

Previous research has shown that, in many contexts, the focus on indicators in research evaluation has had adverse effects, especially in terms of papers published (e.g. Binswanger, 2015). This paper aimed to provide an original analysis of one of these adverse effects, which is to aim to *publish as much as possible*. Our results have shown that, especially for older researchers, the higher the number of papers published throughout their careers, the higher the share of these papers ends up being amongst the top cited papers of their fields. This effect was higher for *Biomedical and health sciences* and for *Social sciences and humanities*, but in all fields the most active group of researcher was also having a higher share of top cited papers. A general exception to this trend was found in academically-younger researchers working in the field of *Life and earth sciences*, where higher scientific output was associated with lower impact than low-to-mid scientific output. Decreasing returns to scale were also more common for more junior researchers than senior ones.

These results conform to the Mertonian theory of cumulative advantages (Merton, 1968): the higher the number of papers an author contributes to, the more he or she gets known and, hence, is likely to attract citations. In Bourdieusian terms, the more an author publishes and accumulates citations in a domain, the more this capital will yield additional papers and citations. The relationship could also be in the other direction, as highly cited authors might have more opportunities to contribute to papers, given the scientific capital they have accumulated. Still, the results show that top cited authors do not only contribute on average to more papers, but also to more *highly cited* papers. On the whole, these results suggest that, at the macro-level, the culture of publishing *as many papers as possible* did not yield to adverse effects in terms of impact, especially for senior researchers. For such researchers, who have had a long period of time to accumulate scientific capital, there can never be *too many papers*.

References

- Abraham, P. (2000). Duplicate and salami publications. *Journal of Postgraduate Medicine*, 46(2), 67.
- Binswanger, M. (2015). How nonsense became excellence: Forcing professors to publish. In Welp, I.M. et al. (eds). *Incentives and Performance* (pp. 19-32). Switzerland: Springer International Publishing.
- Bornmann, L., & Daniel, H. D. (2007). Multiple publication on a single research study: does it pay? The influence of number of research articles on total citation counts in biomedicine. *Journal of the American Society for Information Science and Technology*, 58(8), 1100-1107.
- Bourdieu, P. (2004). *Science of Science and Reflexivity*. Cambridge, UK: Polity Press.
- Butler, L. (2003). Modifying publication practices in response to funding formulas. *Research Evaluation*, 12(1), 39-46.
- Cameron, B. D. (2005). Trends in the usage of ISI bibliometric data: Uses, abuses, and implications. *portal: Libraries and the Academy*, 5(1), 105-125.
- Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, 2(1), 67-90.
- Caron, E., & van Eck, N. J. (2014). Large scale author name disambiguation using rule-based scoring and clustering. In *19th International Conference on Science and Technology Indicators. Context Counts: Pathways to Master Big Data and Little Data* (pp. 79-86). CWTS-Leiden University Leiden.
- Cole, J. R., & Cole, S. (1973). *Social Stratification in Science*. Chicago: University of Chicago Press.
- Costas, R., & Bordons, M. (2007). The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics*, 1(3), 193-203.
- Costas, R., & Bordons, M. (2008). Is g-index better than h-index? An exploratory study at the individual level. *Scientometrics*, 77(2), 267-288.
- CWTS. (2013). *Leiden Ranking 2013 Methodology*. <http://www.leidenranking.com/Content/CWTS%20Leiden%20Ranking%202013.pdf>
- Dahler-Larsen, P. (2011). *The Evaluation Society*. Palo Alto, CA: Stanford University Press.
- Debackere, K., & Glänzel, W. (2004). Using a bibliometric approach to support research policy making: The case of the Flemish BOF-key. *Scientometrics*, 59(2), 253-276.
- Frey, B. S., & Osterloh, M. (2006). *Evaluations: Hidden Costs, Questionable Benefits, and Superior Alternatives*. Institute for Empirical Research in Economics, University of Zurich.
- Fuyuno, I., & Cyranoski, D. (2006). Cash for papers: putting a premium on publication. *Nature*, 441(7095), 792-792.
- Gillespie, R. (1993). *Manufacturing Knowledge: A History of the Hawthorne Experiments*. Cambridge, New York: Cambridge University Press.
- Goodhart, C.A.E. (1975). *Problems of Monetary Management: The U.K. Experience*. Papers in Monetary Economics (Reserve Bank of Australia).
- Hanssen, T. E. S., & Jørgensen, F. (2015). The value of experience in research. *Journal of Informetrics*, 9(1), 16-24.
- Haustein, S., & Larivière, V. (2015). The use of bibliometrics for assessing research: possibilities, limitations and adverse effects. In Lempe, I.M. et al. (eds.) *Incentives and Performance* (pp. 121-139). Switzerland: Springer International Publishing.
- Jefferson, T. (1998). Redundant publication in biomedical sciences: Scientific misconduct or necessity? *Science and Engineering Ethics*, 4(2), 135-140.
- Merton, R. K. (1968). The Matthew effect in science. *Science*, 159(3810), 56-63.
- Sörlin, S. (2007). Funding diversity: performance-based funding regimes as drivers of differentiation in higher education systems. *Higher Education Policy*, 20(4), 413-440.
- Waltman, L., & Van Eck, N. J. (2012). The inconsistency of the h-index. *Journal of the American Society for Information Science and Technology*, 63(2), 406-415.
- Weingart, P. (2005). Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics*, 62(1), 117-131.

Research Assessment and Bibliometrics: Bringing Quality Back in

Michael Ochsner¹ and Sven E. Hug²

¹ ochsner@gess.ethz.ch

ETH Zurich, D-GESS, Mühlegasse 21, 8001 Zürich (Switzerland) and FORS, c/o University of Lausanne, Géopolis, 1015 Lausanne (Switzerland)

² sven.hug@gess.ethz.ch

ETH Zurich, D-GESS, Mühlegasse 21, 8001 Zürich (Switzerland) and University of Zurich, Evaluation Office, Mühlegasse 21, 8001 Zürich (Switzerland)

Introduction

Bibliometric indicators are used to compare research performances and also to assess and evaluate research performance (see, e.g. Gimenez-Toledo et al., 2007; Lane, 2010). However, recently scholars voice protest against bibliometric assessments (see, e.g., Lawrence, 2002; Molinier & Bodenhausen, 2010; Drubin, 2014). The arguments put forward are manifold. For example, the application of the impact factor, which is often used, but not meant, to evaluate individual researchers, is criticized (DORA, 2013). Then, there are myriads of perverse or unintended effects, like focus on high impact journals and mainstream topics, focus on review articles and short communications, strategic behavior, or lack of replication because of the low reputation of replication studies (e.g., Butler, 2007; Lawrence, 2003; Mooneshinghe et al., 2007). Furthermore, scholars from the social sciences and humanities (SSH) criticize that bibliometric indicators cannot capture quality (e.g., Plumpe, 2009).

The authors of this paper were involved in a project to develop quality criteria and indicators for humanities research (see <http://www.psh.ethz.ch/crus>). Here, we argue that while bibliometric indicators and methods are powerful tools to describe research practices and, to some extent, scientific impact, there are some problems when they are readily used as quality indicators in research assessments. We feel that also other disciplines can learn from the critique of humanities scholars on simplistic quantitative assessments and from the findings of the research on quality in the humanities.

Notions of quality

The aim of the project “Developing and Testing Research Quality Criteria in the Humanities” was to find quality criteria and indicators that were at the same time accepted by the humanities scholars and implementable in different linguistic, cultural, and disciplinary settings. Analyzing the humanities scholars’ critique, we found that the development of criteria must take into account the disciplinary research practices, that the measurement must be

transparent and consensual, and that the notions of quality must be made explicit (Hug et al., 2014). We used the Repertory Grid technique to make the notions of quality explicit and base the development of quality criteria on the actual research practices. We found that there are two different conceptions of quality, a more traditional one, which can be described with individual, ground-breaking research that opens up new paradigms, and a more modern conception that can be described as interdisciplinary, project-focused, and public-oriented. Both kind of research can be good as well as bad (Ochsner et al., 2013). Hence, interdisciplinarity, for example, differentiates between two different ways of doing research but is not an indicator of quality (interdisciplinarity can point to good research, when it merges different theories and methods, but it can equally point to bad research that uses interdisciplinarity only for getting funding or for the career). Therefore, notions of quality should be taken into account in research evaluations. They might shed light on gaming strategies as well as on problems with indicators that are not linked to research practices or research quality.

Catalogue of quality criteria

Using the notions of quality, we developed a catalogue of quality criteria that are linked to the research practices in the humanities. Humanities scholars then rated these criteria as well as indicators measuring those criteria. We found that a broad range of quality criteria and aspects must be taken into account to adequately assess research quality (Hug et al., 2013) and that only about 3% to 32% of the scholars’ notions of quality can be quantified adequately, depending on the discipline. Furthermore, we found that there is a mismatch between the quality criteria put forward by the scholars and the quality criteria used in evaluation procedures (Ochsner et al., 2012). Hence, current evaluation procedures do not measure research quality in the humanities adequately. This does not mean that the existing evaluation procedures and criteria are useless (e.g., societal impact is not necessarily linked to research quality but is a legitimate criterion in evaluations), but it shows that

a very important dimension of research assessment is not reflected adequately: quality of research.

The humanities, so what?!

Our research bases on the humanities. What is the relevance of this research to the rest of academia? First, we argue that humanities scholars, while not specialised in quantification, are experts in critical thinking. Hence, their critique of evaluation procedures often points to the consequences of the instruments on research practices. This is what increasingly also happens in the natural sciences (e.g., DORA, 2013; Drubin, 2014) because some perverse effects start to become apparent. Hence, a focus on research practices in assessments could help minimise negative impact of indicators. Second, when we presented the criteria at conferences and workshops, also natural scientists were present. They surprisingly often said that the criteria we presented made also sense to them with a few exceptions. Hence, what could be learned from the case of the humanities would be the following: base evaluation procedures on research practices; be aware that the indicators used will affect the research practices; formulate quality criteria in a way that makes sense to the scholars; involve as many stakeholders as possible in the definition of quality criteria.

Bringing quality back in

While the bibliometric community is well aware of the possible drawbacks of bibliometric indicators, the most common reaction by the research evaluation community is to look for other sources of the same kind of indicators and altmetrics. We think that the problem is not a technical one but a conceptual. At the beginning of any research evaluation and science policy should be a reflection on the goals. Do we want scholars to use most of their time to feed Twitter, comment on Research Gate, or 'pimp' their statistics in Google Scholar? We think that research evaluation should bring quality back in. Evaluation and assessments should not solely *judge* the merits of scholars but help them to *enhance* their impact by fostering research quality. Hence, bibliometrics and altmetrics are powerful instruments to describe certain impacts, visibility, networks etc. But research assessments should also make clear statements about other aspects of research quality. Therefore, the disciplinary community should have a say in what criteria are applied in their assessments. New ideas of research evaluation based on research practices should lead scientific discussion much more than technical issues vaguely related to research quality.

Acknowledgments

This paper is based on work that was supported by the Rectors' Conference of the Swiss Universities

(CRUS) within the framework of the SUK B-05 Innovation and Cooperation Project "Mesurer les performances de la recherche". Matching funds were provided by the University of Zurich.

References

- Butler, L. (2007). Assessing university research: a plea for a balanced approach. *Science and Public Policy*, 34(8), 565-574.
- DORA. (2013). San Francisco Declaration on Research Assessment. <http://am.ascb.org/dora/>
- Drubin, D. (2014). Time to discard the metric that decides how science is rated. *The Conversation*. <http://theconversation.com/time-to-discard-the-metric-that-decides-how-science-is-rated-27733>
- Gimenez-Toledo, E., Roman-Roman, A. and Alcain-Partearroyo, M. D. (2007). From Experimentation to Coordination in the Evaluation of Spanish Scientific Journals in the Humanities and Social Sciences. *Research Evaluation*, 16(2), 137-48.
- Hug, S. E., Ochsner, M., & Daniel, H.-D. (2013). Criteria for assessing research quality in the humanities: a Delphi study among scholars of English literature, German literature and art history. *Research Evaluation*, 22(5), 369-383.
- Hug, S. E., Ochsner, M., & Daniel, H.-D. (2014). A framework to explore and develop criteria for assessing research quality in the humanities. *International Journal for Education Law and Policy*, 10(1), 55-64.
- Lane, J. (2010). Let's Make Science Metrics More Scientific. *Nature*, 464(25), 488-489.
- Lawrence, P.A. (2002). Rank injustice. The misallocation of credit is endemic in science. *Nature*, 415, 835-836.
- Lawrence, P.A. (2003). The politics of publication. Authors, reviewers and editors must act to protect the quality of research. *Nature*, 422, 259-261.
- Molinie, A. & Bodenhausen, G. (2010). Bibliometrics as Weapons of Mass Citation. *Chimia* 64(1-2), 78-89.
- Mooneshinghe, R., Khoury, M. J., & Janssens, A. C. J. W. (2007). Most published research findings are false - but a little replication goes a long way. *PLOS Medicine*, 4(2), e28.
- Ochsner, M., Hug, S. E., & Daniel, H.-D. (2012). Indicators for research quality in the humanities: opportunities and limitations. *Bibliometrie - Praxis und Forschung*, 1(4).
- Ochsner, M., Hug, S. E., & Daniel, H.-D. (2013). Four types of research in the humanities: setting the stage for research quality criteria in the humanities. *Research Evaluation*, 22, 79-92.
- Plumpe, W. (2009). Stellungnahme zum Rating des Wissenschaftsrates aus Sicht des Historikerverbandes. In C. Prinz & R. Hohls (Eds.), *Qualitätsmessung, Evaluation, Forschungsrating. Risiken und Chancen für die Geschichtswissenschaft?* (pp. 121-126). Historisches Forum. Berlin: Clio-online.

Under-reporting research relevant to local needs in the global south. Database biases in the representation of knowledge on rice

Ismael Rafols,^{1,2} Tommaso Ciarli² and Diego Chavarro²

¹*i.rafols@ingenio.upv.es*

Ingenio (CSIC-UPV), Universitat Politècnica de València, València, Spain,
SPRU (Science and Technology Policy Research), University of Sussex, Brighton, UK,
and Observatoire des Sciences et Techniques (HCERES-OST), Paris, France

²*t.ciarli@sussex.ac.uk, diego.chavarro@sussex.ac.uk*

SPRU, Science Policy Research Unit, University of Sussex, Brighton, UK

Introduction

There is an increasing demand for science to help in addressing grand challenges or societal problems, such as tackling obesity, climate change or pandemics. In this context, it becomes important to understand what different sciences can offer to tackle these problems, and towards which directions scientific research should be developed. A useful starting point is to investigate what is the existing science supply, and which research options are better aligned to address grand challenges and societal demands (Sarewitz & Pielke, 2007). In order to map the science supply, we need a representation of the knowledge on research topics relevant for a problem.

Bibliometrics can provide very helpful tools for developing knowledge representations. However, these representations are highly dependent on the data and methods used. As a result, bibliometric tools or indicators often reproduce the biases in the data collection and treatment. For example, it has been shown that conventional bibliometric analyses are biased against non-English languages (Van Leeuwen et al., 2001), developing countries (Velho & Krige, 1986), applied science (Van Eck et al., 2013), the social sciences and humanities (Martin et al., 2010) and interdisciplinary research (Rafols et al., 2012). The aim of this paper is to investigate the biases introduced by available databases in the representation of research topics.

In a previous study on rice research, we showed that the bibliographic database CAB Abstracts (CABI) – which is focussed on agriculture and global health – has a larger coverage of rice research for most low income countries than Web of Science (WoS) or Scopus (Ciarli, Rafols & Llopis, 2014). For example, India has twice the number of publications in CABI on rice compared to Scopus and about 4 times those in WoS. In this study, we present evidence that shows that this unequal coverage distorts significantly the knowledge representation of rice research, globally and for different countries. Such bias may have

policy effects, in particular for a societal issue such as rice production.

As shown in Figure 1, we find that the journal coverage of the bibliometric databases WoS and Scopus under-represent some of the more application oriented topics (namely: i) production, productivity and plant nutrition (top left); ii) plant characteristics (top center); and iii) diseases, pests and plant protection (center).

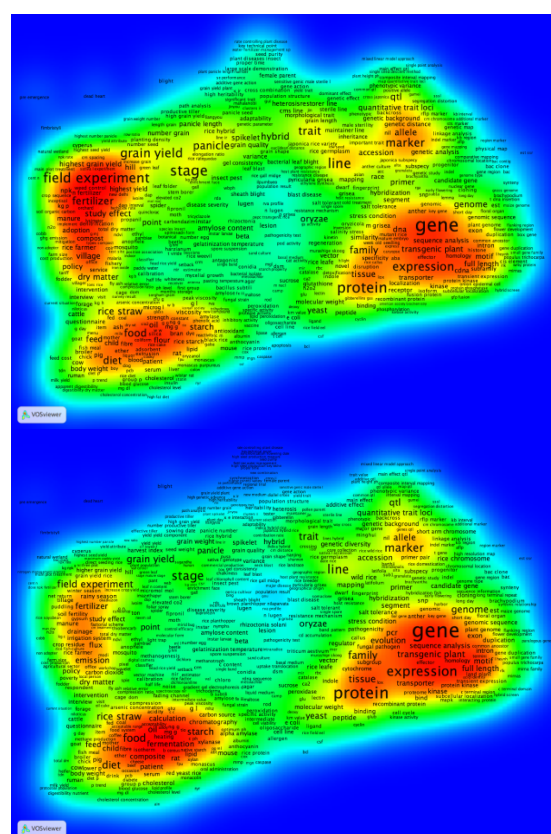
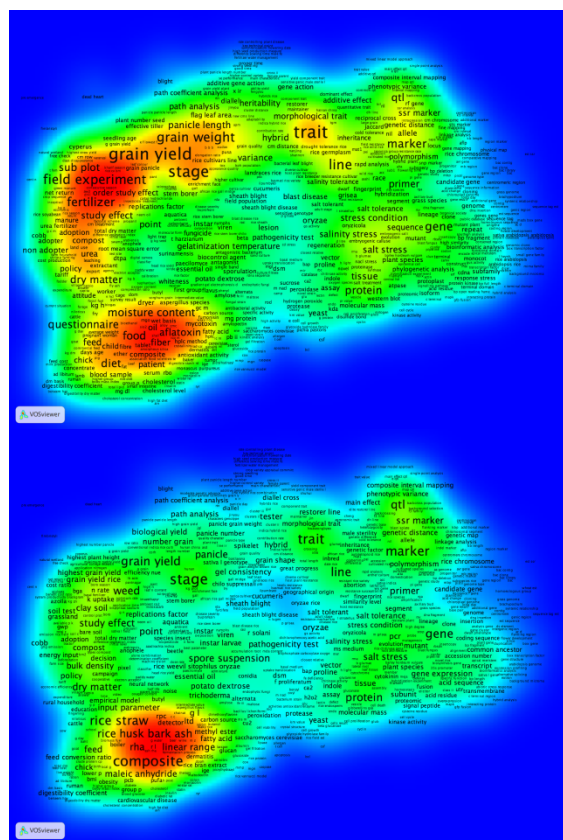


Figure 1. Publication density for rice research in CABI (top) and in WoS (bottom). The top left and top right areas under-report in WoS are related to production and seed characteristics.

Given that these are issues relevant to small farmers, producing for the local market, and with no access to the seeds developed with molecular biology techniques (GM – bottom left), we pose the

Figure 2 illustrates that under-representation of research on production, pest and seed characteristics is particularly acute in some countries with molecular biology research (related to GM), but with a focus on research to address food security and local farming needs (in this case Iran). Rice research in these countries tends to be more focused on increasing crop yield, precisely the topic under-represented in WoS and Scopus.



Conclusions

farmers and consumers local needs. This study does not answer this question, but it shows that it is a meaningful and important issue for bibliometrics to address: bibliometric exercise that use dominant databases may have a negative effect on policies relevant to important social issues, particularly in developing countries.

Information on methods and data

Acknowledgments

We acknowledge support from the EU (Marie Curie Integration fellowship to IR), the UK ESRC (RES-360-25-0076) and the US NSF (Award #1064146). The findings and observations contained in this paper are those of the authors and do not necessarily reflect the views of the funders.

References

- Ciarli, T., et al. (2014). The under-representation of developing countries in the main bibliometric databases. *Proceedings of the S&T Indicators Conference* (97–105). Leiden.
- Martin, B. R., Tang, P., Morgan, M., & al. (2010). *Towards a Bibliometric Database for the Social Sciences and Humanities – A European Scoping Project* (A report for DFG, ESRC, AHRC, NWO, ANR and ESF). Brighton, UK: SPRU.
- Rafols, I., et al. (2012). How journal rankings can suppress interdisciplinarity. *Research Policy*, 41(7), 1262–1282.
- Sarewitz, D., & Pielke, R. A. (2007). The neglected heart of science policy: reconciling supply of and demand for science. *Environmental Science & Policy*, 10(1), 5–16.
- Van Eck, N. J. et al. (2013). Citation Analysis May Underestimate the Impact of Clinical Research as Compared to Basic Research. *PLoS ONE*, 8(4), e62395. doi:10.1371/journal.pone.0062395
- Van Leeuwen, T. N., et al. (2001). Language biases in the coverage of the Science Citation Index. *Scientometrics*, 51(1), 335–346.
- Velho, L., & Krige, J. (1984). Publication and citation practices of Brazilian agricultural scientists. *Social Studies of Science*, 14(1), 45–62.

Network DEA approach for measuring the efficiency of University-Industry Collaboration Innovation: Evidence from China

Yu Yu¹, Qinfen Shi² and Jie Wu³

¹ *yuyu0801@139.com*

Hohai University, Business School, No.8 Focheng Road West, 211100 Nanjing (China)

² *shiqf@njupt.edu.cn*

Nanjing University of Posts and Telecommunications, School of Management, No.9 Wenyuan Road, 210023 Nanjing (China)

³ *0511wujie@163.com*

Jiangsu University of Science and Technology, School of Economics and Management, No.2 Mengxi Road, 212003 Zhenjiang (China)

Introduction

Collaborative innovation is a trans-disciplinary approach for developing the wholeness synergy to improve the competitiveness of an organization through holistic, competitive and complementary interactions between and among innovation participants in a specific environment (Bommert, 2010; Swink, 2006). The collaborative innovation system essentially consists of three sectors: industry, universities, and the government, with each sector interacting with the others, while at the same time playing its own role. Collaborative innovation system is a complex conglomerate of interacting independent parties. The network of institutional relations among universities, industries, and governmental agencies has been considered as a Triple Helix (TH). Collaborative innovation system (CIS) is based on a multi-input, multi-output transformation relation. It is an important issue to investigate the performance related to the transformation process of limited innovation resources for improving collaborative innovative outputs. Previous studies have been done to evaluate the performance of collaborative innovation. However, those studies failed to consider the complexity of the collaborative innovation system. Data envelopment analysis (DEA) is a method for measuring the efficiency of peer decision making units (DMUs). Recently network DEA models been developed to examine the efficiency of DMUs with internal structures. The internal network structures range from a simple two-stage process to a complex system where multiple divisions are linked together with intermediate measures. In this study, we propose a network DEA with parallel production systems to measure the efficiency of University-Industry Collaborative Innovation. The purpose of the present study is to construct a complete measurement framework characterizing the CIS' production framework from original S&T

investment to final outputs, and measure the CIS' process-oriented technical efficiency, which is implemented in China's context. It is hoped that this study will benefit China's collaborative innovation policy-making.

Network DEA model

We propose a network DEA with parallel production systems in this section. Assume that there are n DMUs, and each DMU has two *sub-DMUs*. Figure 1 depicts the visual structure of the DEA model.

The part of inputs is consumed by SDMU1 and SDMU2 together, and part of DMU output is co-produced by SDMU1 and SDMU2. Besides, some inputs and outputs are consumed or produced by SDMU1 or SDMU2 alone. Variables are defined as follows: $X_1 = (x_{1j}^1, K, x_{mj}^1)$ represent m separate inputs which are consumed by SDMU1; $X_2 = (x_{1j}^2, K, x_{hj}^2)$ represent h separate inputs which are consumed by SDMU2; $X_s = (x_{1j}^s, K, x_{lj}^s)$ represent l inputs consumed by SDMU1 and SDMU2 together. The vector of $Y_1 = (y_{1j}^1, K, y_{sj}^1)$ are s outputs produced by SDMU1; the vector of $Y_2 = (y_{1j}^2, K, y_{tj}^2)$ are t outputs produced by SDMU2; the vector of $Y_s = (y_{1j}^s, K, y_{uj}^s)$ are u outputs produced by SDMU1 and SDMU2 together.

For analytical tractability, we use $X_{s1} = (x_{1j}^{s1}, K, x_{lj}^{s1})$, $X_{s2} = (x_{1j}^{s2}, K, x_{lj}^{s2})$, $Y_{s1} = (y_{1j}^{s1}, K, y_{uj}^{s1})$ and $Y_{s2} = (y_{1j}^{s2}, K, y_{uj}^{s2})$ to represent the shared inputs and outputs of SDMU1 and SDMU2 in each subsystem, and $X_s = X_{s1} + X_{s2}$, $Y_s = Y_{s1} + Y_{s2}$.

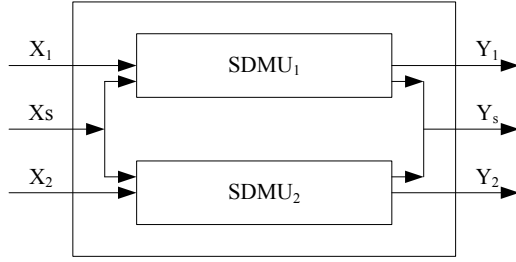


Figure 1. Parallel system structure.

In this study, we choose new product sales as independent output in Industry sub-system, the number of universities' published papers as independent output in universities sub-system. Patent applications in IU collaboration innovation system mainly come from both industry and universities subsystems; therefore the number of patent applications is seen as a shared output in the system.

According to DEA parallel production system efficiency evaluation model proposed by Kao (2009), parallel production system efficiency of the DMU under constant returns to scale (CRS) can be represented as follows:

$$\begin{aligned} \bar{\theta}_{CRS}^* &= \min \theta \\ s.t. \\ \sum_{k=1}^2 \sum_{j=1}^n \lambda_j^k y_{rj}^{sk} &\geq y_{ro}^s \quad r = 1, K, u \\ \sum_{j=1}^n \lambda_j^1 y_{rj}^1 &\geq y_{ro}^1 \quad r = 1, K, s \\ \sum_{j=1}^n \lambda_j^2 y_{rj}^2 &\geq y_{ro}^2 \quad r = 1, K, t \\ \sum_{k=1}^2 \sum_{j=1}^n \lambda_j^k x_{ij}^{sk} &\leq \theta x_{io}^s \quad i = 1, K, l \\ \sum_{j=1}^n \lambda_j^1 x_{ij}^1 &\leq \theta x_{io}^1 \quad i = 1, K, m \\ \sum_{j=1}^n \lambda_j^2 x_{ij}^2 &\leq \theta x_{io}^2 \quad i = 1, K, h \\ \sum_{j=1}^n \lambda_j^1 &= \sum_{j=1}^n \lambda_j^2 \\ \lambda_j^k &\geq 0 \quad k = 1, 2; j = 1, K, n \end{aligned}$$

The main data in this paper are all selected in the "China Statistical Yearbook of Science and Technology". Considering the time lag in innovation activities, we select the data in 2009 as input data and the data in 2010 as output data in this paper. This study excludes all provinces that have missing data. Finally, this study evaluates 30 observations of Chinese provinces.

Table 1 summarizes three efficiency scores under constant returns to scale (CRS), variable returns to scale (VRS) and non-increasing returns to scale (NIRS).

Table 1. Three Efficiencies of Chinese provinces.

Province	$\bar{\theta}_{CRS}^*$	$\bar{\theta}_{NIRS}^*$	$\bar{\theta}_{VRS}^*$
Beijing	0.5903	1.0000	1.0000
Tianjin	0.9412	1.0000	1.0000
Hebei	0.6656	0.6656	0.6692
Shanxi	0.3089	0.3089	0.3189
Inner Mongolia	0.4715	0.4715	0.4974
Liaoning	0.4605	0.4605	0.4636
Jilin	1.0000	1.0000	1.0000
Heilongjiang	0.3869	0.3869	0.3882
Shanghai	0.8232	1.0000	1.0000
Jiangsu	0.8229	1.0000	1.0000
Zhejiang	0.8769	0.8791	0.8791
Anhui	0.6534	0.6546	0.6546
Fujian	0.5968	0.5968	0.6002
Jiangxi	0.5474	0.5474	0.5491
Shandong	0.6453	1.0000	1.0000
Henan	1.0000	1.0000	1.0000
Hubei	0.6291	0.8497	0.8497
Hunan	0.6651	0.6667	0.6667
Guangdong	0.8773	1.0000	1.0000
Guangxi	0.7016	0.7016	0.7095
Hainan	0.9648	0.9648	1.0000
Chongqing	0.9698	0.9903	0.9903
Sichuan	0.4845	0.5530	0.5530
Guizhou	0.6488	0.6488	0.6661
Yunnan	0.5810	0.5810	0.6081
Shaanxi	0.6860	0.6860	0.6861
Gansu	0.8782	0.8782	0.8828
Qinghai	0.3233	0.3233	0.8972
Ningxia	0.5769	0.5769	0.6545
Xinjiang	0.7036	0.7036	0.7416

Results

The average efficiency under constant returns to scale of University- Industry collaborative innovation in China is 0.7642. However, the efficiencies of some provinces are less than the average efficiency. By the view of economic region, the efficiencies of UI collaborative innovation in eastern, northern and southern coastal China are higher than other areas in China.

Acknowledgments

The work is supported by National Natural Science Foundation of China (No. 71471091, 71271119).

References

- Bommert, B. (2010). Collaborative innovation in the public sector. *International Public Management Review*, 11(1), 15-33
- Kao, C. (2009). Efficiency measurement for parallel production systems. *European Journal of Operational Research*, 196(3), 1107-1112.
- Swink, M. (2006). Building collaborative innovation capability. *Research-technology Management*, 49(2), 37-47.

Promotions, Tenures, and Publication Behaviours: Serbian Example

Dejan Pajić and Tanja Jevremov

dpajic@ff.uns.ac.rs; tanja.jevremov@uns.ac.rs

University of Novi Sad, Faculty of Philosophy, Department of Psychology,
Dr Zorana Đinđića 2, 21000 Novi Sad (Serbia)

Introduction

Bibliometric indicators became a common tool for evaluating universities (Geuna & Martin, 2003). Furthermore, individual academics and researchers are also evaluated, promoted, and tenured based on their productivity, particularly the one visible in international databases such as the Web of Science (WoS). This methodology is widely accepted even in non-English speaking countries (Pajić, 2014).

Growing emphasis on bibliometric indicators is followed by a continuing debate on their suitability for the evaluation in social sciences and humanities (SS&H) (Nederhof, 2006). Secondary importance of journals and the prevalence of monographs are usually identified as the key features of "publication behaviour" in SS&H (Hicks, 2012). Economics and psychology are often considered to be more similar to sciences (Engels, Ossenblok, & Spruyt, 2012).

This paper presents initial results on the scientific productivity of professors promoted and tenured at the University of Novi Sad (UNS). The main goal was to analyse publication patterns in SS&H and their implications for the evaluation of individuals.

Data and method

UNS is the second largest state university in Serbia. It consists of 14 faculties and 2 research institutes. Presented analysis was focused on the production of professors promoted or tenured in 2009-2013 at 6 UNS faculties in SS&H. Data were taken from the reports publicly available on the UNS website¹. Each report contained bibliography provided by the candidate and was verified by the corresponding committee of at least three members.

The sample included 297 professors in language and literature (99), education (62), economics (32), psychology (27), law (26), history (19), sociology (12), philosophy (10), and science (e.g. professors of chemistry at teachers colleges) (10). The total of 9007 publications were extracted and categorized according to the origin (national, international), and type (books, journal articles, proceedings, other). In order to balance the differences in the publication counts among the researchers of different academic rank, only publications from the last promotion period of 5 years were taken into account.

¹ <http://www.uns.ac.rs/sr/izborZvanje/bilteni.html> Reports were removed during the preparation of this paper and are no longer available online, but are available from the authors.

Since this is a preliminary analysis, it was mostly based on descriptive statistics. Because of skewed distributions, non-parametric tests were used to test the basic differences among disciplines.

Results and discussion

Kruskal-Wallis test indicates significant differences in scientific productivity among researchers from different fields: $H(8, 297) = 22.99, p < .01$ (Figure 1). It is difficult to draw a solid conclusion, mainly because of highly skewed distributions and large individual differences, but clearly psychology and sciences have the highest median values, while the lowest scientific activity is that of the researchers in the field of law. The most pronounced individual differences were observed in the fields of language and literature, and educational sciences.

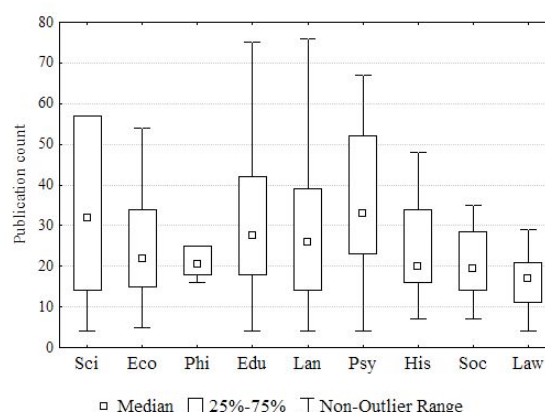


Figure 1. Differences in scientific productivity among researchers in nine scientific fields.

Distributions of the major types of publications among scientific fields differ significantly: $\chi^2(16, 8492), p < .01$ (Figure 2). The share of articles is somewhat unusually high in humanities, and ranges around 40% in all fields. Contrary to usual beliefs, psychology and sciences have the lowest proportion of journal articles within the total number of publications. On the other hand, the highest proportion was detected in the field of law where journal articles account for almost 2/3 of all publications. However, the list of the most frequent journal titles revealed that more than half of the articles were from a journal published by the same faculty where the candidates were promoted or tenured.

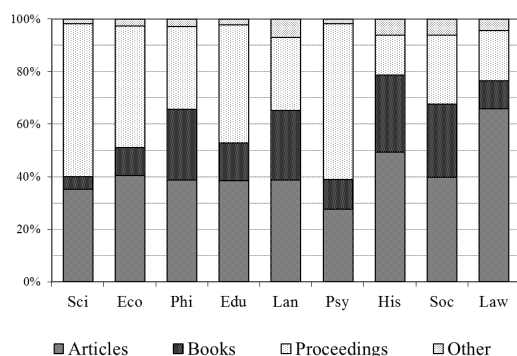


Figure 2. Proportions of different types of publications in nine scientific fields.

Our results have confirmed the importance of book chapters and monographs in humanities, although this type of publication is not predominant in any of the fields. Conference abstracts and proceedings are the most frequent type of publication in four out of nine analyzed fields.

Figure 3 shows the proportions of (inter)national publications across scientific fields. The strongest focus on international sources is noticeable in the sciences, and the lowest in history, sociology, and law. The results that are not in line with the usual beliefs are rather nationally oriented publication behavior of Serbian psychologists, and a relatively high ratio of international sources in philosophy.

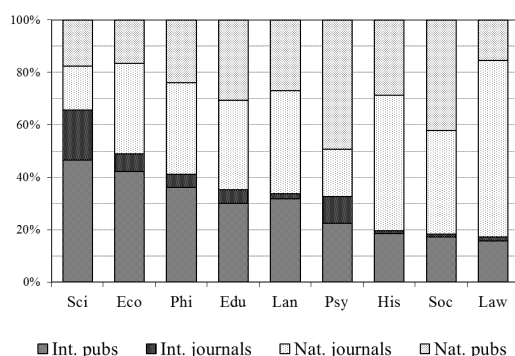


Figure 3. Proportions of national and international publications in nine scientific fields.

Professors at the faculties in Serbia are required to have one to three papers published in WoS journals prior to promotion or tenure. Table 1 shows the list of the 15 most common (allegedly) WoS journals reported in 297 reports. The majority of journals are actually national or regional WoS journals with the rather low impact factor values (IF). The disturbing fact is that several professors were promoted based on their articles published in journals of dubious quality, those that were dropped from WoS because of academic malpractice (e.g. *HealthMED*, *TTEM*, *Metalurgia Int*) or were never indexed by WoS nor any major international bibliographic database (e.g. *Brit Amer Stud*). In addition, 12 other journals were falsely reported as top ranked WoS titles.

Table 1. Most common (allegedly) WoS journals listed in 297 promotion and tenure reports.

Journal title	%	Country	IF
Psihologija	17.50	SRB	0.188
TTEM	5.83	B&H	drop.
HeathMED	5.13	B&H	drop.
Croat J Educ	3.03	CRO	0.034
Roman J Eng Stud	2.30	ROM	-
Med Sport	2.30	ITA	0.125
Vojnosan pregl	2.10	SRB	0.269
New Edu Rev	1.63	POL	drop.
Filoz istraživanja	1.63	CRO	AHCI
Brit Amer Stud	1.40	ROM	-
Panoeconomicus	1.16	SRB	0.778
Riječ	1.16	CRO	-
Didactica Slov	0.93	SLO	drop.
ICCCC	0.93	ROM	0.694
Metalurgia Int	0.93	ROM	drop.

drop. - dropped from WoS

Conclusion

Our results have shown that SS&H are clearly more nationally oriented compared to sciences. However, journals as knowledge dissemination channels seem to be equally important across all fields. Apart from the conference proceedings, journal articles are the most common type of publications. It's obvious that the current promotion and tenure rules affect the professors' publication behaviour. Such patterns are not determined simply by the characteristics of a discipline, but in some cases by the ease of access to particular sources, e.g. journals having a rather lenient editorial policy.

Science policy institutions should be aware that the evaluation is a dynamic process that must combine both the rules and the means to assess the effects of those rules and to monitor their implementation.

References

- Engels, T.C.E., Ossenblok, T.L.B., & Spruyt, E.H.J. (2012). Changing publication patterns in the Social Sciences and Humanities, 2000-2009. *Scientometrics*, 93(2), 373-390.
- Geuna, A., & Martin, B R. (2003). University Research Evaluation and Funding: An International Comparison. *Minerva*, 41(4), 277-304.
- Hicks, D. (2012). One size doesn't fit all: on the co-evolution of national evaluation systems and social science publishing. *Confero*, 1(1), 67-99.
- Nederhof, A.J. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review. *Scientometrics*, 66(1), 81-100.
- Pajić, D. (2015). Globalization of the social sciences in Eastern Europe: genuine breakthrough or a slippery slope of the research evaluation practice? *Scientometrics*, 120(3), 2131-2150.

The Serbian Citation Index: Contest and Collapse

Dejan Pajić

dpajic@ff.uns.ac.rs

University of Novi Sad, Faculty of Philosophy, Department of Psychology,
Dr Zorana Đinđića 2, 21000 Novi Sad (Serbia)

The Past

Ten years ago, a poster titled *The Serbian Citation Index: context and content* was presented at the ISSI conference held in Stockholm (Šipka, 2005). *Serbian Citation Index (SCIndeks)* was at the time a pioneering effort to build a comprehensive, open access citation index of Serbian scientific journals with three missions: local *dissemination* of research findings in the open access mode, global *promotion* of the Serbian science, and objective *evaluation* of national journals, institutions, and researchers. Started as an ambitious project of the group of enthusiasts and volunteers in 1990s, SCIndeks has become truly embraced nationally during the 2000s. In the period when Serbia was represented in the Web of Science (WoS) with only three journals, SCIndeks was recognized as a tool to enhance the public accountability, visibility, and quality of local journals. *Centre for Evaluation in Education and Science (CEES)*, SCIndeks developer and publisher, started receiving full financial support from the Serbian Ministry of Science (SMS), both for the maintenance of SCIndeks and for publishing the *Journal Bibliometric Report (JBR)*. The report is published annually and contains the national impact factor and almost 20 other bibliometric indicators for over 300 journals covered by SCIndeks. JBR is used for journal rankings and, indirectly, as a data source for the evaluation of individual researchers, their promotions, and tenures.

The Contest

The role and importance of a national citation index cannot be evaluated outside the global scientific information market. The first test for SCIndeks was the recognition and perception of Serbian journals by the major international database providers. After Elsevier's Scopus and Google's Scholar appeared in 2004, Thomson Reuters' indexing policy has also changed radically. The question was whether the CEES efforts to improve the visibility and quality of local journals would result in increased number of titles accepted for indexing in WoS and Scopus. Figure 1 shows the number of journals published in Serbia and three neighbouring countries indexed in WoS and Scopus. All countries have managed to improve their visibility in international databases, but the Serbian progress is only slightly ahead of Bulgarian and far behind Romanian and Croatian. Neither Bulgaria nor Romania has national citation

index or a repository of national journals. On the other hand, Croatian journals are presented in the *Portal of Scientific Journals of Croatia* and the *Croatian Scientific Bibliography*, both funded by the government, but having limited functionality compared to SCIndeks, especially regarding the support for journal editors, evaluators, and science policy institutions. It seems that the mission to promote journals through SCIndeks has failed or at least has not succeeded in lowering a potential bias in inclusion policies of the major database providers.

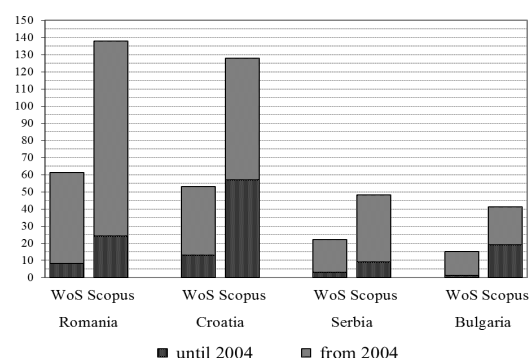


Figure 1. Growth in the number of WoS and Scopus journals published in Serbia and three neighbouring countries.

Another, and perhaps the more important contest, was carried out at the local (political) level. Every assessment brings the risk of conflict of interest. If such an assessment influences the allocation of funds and promotion and tenure decisions, the risk is even higher. Although the government supported CEES financially, it did not fully uphold the practical implementation of CEES reports on the quality of national journals (Šipka, 2014). Journal rankings based on impact measures and SCIndeks data were often altered by the ministerial committees in order to favour the very journals whose editors were members of those committees. In some cases, worst ranked national journals were given the status of international ones. At the level of individuals, it would mean that a candidate for promotion would earn points sufficient for a position of assistant professor by publishing two articles in a bottom-ranked local journal or a journal that was not even accepted for indexing in the national citation index.

The Collapse

In 2014, SMS has ceased to finance both the JBR and SCIndeks. In 2015, the effects of that decision have become visible in the form of significantly reduced SCIndeks coverage. A large amount of data were taken offline and became inaccessible to the users of SCIndeks and other web services, such as Google Scholar. Table 1 shows the amount of this "information market disturbance".

Table 1. SCIndeks data available online before and after the cut of funding.

<i>No. of</i>	<i>Apr. 2008</i>	<i>Apr. 2014</i>	<i>Apr. 2015</i>
journals	357	411	56
abstracts	82.876	151.027	19.900
full texts	23.421	58.068	12.172
references	917.567	2.078.642	335.344

As a response to the CEES' "strategic move", SMS has decided to continue using SCIndeks data for evaluation purposes and to finance JBR after all. However, all journals are now required to pay the indexing fees, including some additional costs for options like the full-text availability, cited reference search and cross-linking within SCIndeks. In short, a communication failure between CEES and SMS anticipates the start of a "natural selection" process for the majority of Serbian academic journals and the collapse of the open science idea in Serbia.

One aspect of this collapse is the fact that tens of thousands of papers written by the authors from Serbia are no longer available online and that additional costs are required for them to reappear. Another equally relevant issue is the profile of journals currently accessible through (what was) the national citation index. All of those journals are willing (or able) to pay the indexing fees, but just a few of them were previously classified as leading national journals. An example of this obvious compromise is the fact that although the diversity of affiliations within journal issues was strongly encouraged by both the national regulations and earlier SCIndeks inclusion guidelines, CEES indexes several journals with the majority of papers written by the authors affiliated with the journal's publishing institution.

The Future

Under the current circumstances, SCIndeks can no longer be considered to be the national citation index. The question is who should be concerned with the fact that it has become a mere commercial product with the special status at SMS. The state is surely a loser in this scenario being unable to claim and protect at least the metadata whose production it financed for several years. As for the Serbian scientific community, its future reactions are maybe not that hard to predict. A certain segment of this community has already expressed their opinion on

this matter through the acts of various interest groups opposing the implementation of evaluation methodology based on SCIndeks data. On the other hand, an increasing number of researchers from Serbia are shifting the focus towards international journals, both when publishing and citing journal articles (Pajić & Jevremov, 2014). The evaluation of national science is hence being either spurned or entrusted to the international publishers and their reviewers. In this context, national citation index is becoming a costly repository whose functionalities will not be missed much by researchers or journal editors. More than 300 Serbian journals are now available online and none of them relies solely on SCIndeks when it comes to the visibility. Although some editors are satisfied with the combination of journal's personal website and free Google Scholar services, the growing number of Serbian journals are also being available through other databases and repositories, such as the *Directory of Open Access Journals*, *ERIH PLUS* or *EBSCO* databases. What was conceived as a joint effort to truly promote Serbian science has turned into an "every man for himself" strategy ten years after.

Conclusion

The basic idea of a national citation index was fully justified in the period of domination of Thomson Reuters' citation indices. But this domination is not nearly as strong as it was before, mainly due to the emergence of Scopus and Scholar. We can consider *SciELO* (now hosted by WoS) as an example of a successfully realized "peripheral" citation index. If this was achieved by covering some 1,200 journals from 12 different countries, then SCIndeks and its 400+ journals tell us how justified is the idea of a national citation index and how ambitious it should be. SCIndeks and its fate is the fate of any self-sufficient and rigid science policy institution, but also the fate of any scientific community that is simply too confined and too small. Too small to neglect the inevitable globalization of science, too small to rely on the integrity of its own members to ensure the quality control, and finally too small to satisfy its own ambitions.

References

- Pajić, D., & Jevremov, T. (2014). Globally national - locally international: Bibliometric analysis of a SEE psychology journal. *Psihologija*, 47(2), 263–277, doi:10.2298/PSI1402263P
- Šipka, P. (2005). The Serbian citation index: context and content. In *Proc. of ISSI 2005, Stockholm, Sweden, July 24-28, 2005* (pp. 710–711).
- Šipka, P. (2014). Methods of evaluation of scientific journals: use and abuse [in Serbian]. In Lj. Vučković-Dekić & N. Arsenijević (Eds.), *Vrednovanje nauke i naučnika* (pp. 9–30). Kragujevac: Fakultet medicinskih nauka & Beograd: Akademija medicinskih nauka.

Selecting Researchers with a Not Very Long Career - The Role of Bibliometrics

Elizabeth S. Vieira¹ and José A. N.F. Gomes²

¹ elizabeth.vieira@fc.up.pt, ² jfgomes@fc.up.pt

REQUIMTE/Departament of Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

Introduction

The scientific community has developed many institutionalized forms of evaluation where peer review has an important role, but recently, bibliometric methods have been gaining some acceptability to assess the scientific performance.

The two techniques have been related to one another in different ways: 1) bibliometric methods have been used to analyze the peer review processes (Moed, 2005, chapters 19 and 20); 2) the peer review process uses bibliometric parameters as an auxiliary instrument (Moed, 2005 chapter 18, p. 233-234); and 3) peer reviewers are called in to validate and correct the results of some bibliometric process (e.g. Norris & Oppenheim, 2003; Rinia, van Leeuwen, van Vuren, & van Raan, 1998). There are some national scientific systems that use bibliometric techniques or a mix of bibliometric techniques and peer review to decide the allocation of funding (e.g. Excellence in Research for Australia (ERA); *Valutazione della Qualità della Ricerca* (VQR)). Taking into account the advantages and limitations of bibliometric techniques and the intensive use, recently, there is a growing interest in its potential in helping peers to prepare the final decisions and therefore several studies have been made on the subject (e.g. Vieira, Cabral, & Gomes, 2014a, 2014b, Bornmann & Leydesdorff, 2013). In this study, we exploit the usability of bibliometrics as support tool this time in selecting candidates that had been awarded their PhD's more than 6 and less than 12 years ago and had worked as independent researchers for less than 6 years. We deem this study important as: (1) there is a growing use of bibliometric indicators and it is important to know their caveats and strong points at the different levels; and (2) the use of bibliometric indicators is more controversial when applied to individual researchers, especially at initial steps of their careers.

Methodology

This study considers the applicants to the development grants of the opening *Investigador FCT* carried out in Portugal since 2012. The publications indexed in the Web of Science Core Collection of the 120 applicants from the Engineering and Technology (28), Natural Sciences

(23), Exact Sciences (48) and Medical and Health Sciences (21) were used to calculate a set of bibliometric indicators that are intended to describe the scientific performance. Bibliometric techniques are not used in a formal way in the opening. However, we are looking for indicators that may be implicit in peer judgments. A set of 17 indicators was determined: *TD* (number of documents); *TDC* (number of cited documents); *NDF* (number of documents after fractionation by the total number of authors); *PA* (% of articles); *PP* (% of proceedings papers); *PR* (% of reviews); *PAP* (% of documents as articles and proceedings papers simultaneously); *PDAC* (% of documents as corresponding author); *h* index, *h_{nf}* index (Vieira & Gomes, 2011); *SNIP_m* (median of all the *SNIPs* of the journals where the applicant has published, Moed, 2010); *SJR_m* (median value as in the *SNIP_m*, Gonzalez-Pereira, Guerrero-Bote, & Moya-Anegón, 2010); *PTDIF* (% of documents published in journals with Impact Factor- *IF*); *PQI* (% of documents published in journals in the first quartile in its scientific domain, according to the *IF*); *HCD* (% of documents highly cited in the top 10%); *NI* (average number of citations per document after normalization); *DIC* (% of documents with international collaboration). There is a huge number of bibliometric indicators and we tried to select those that describe the several dimensions of the scientific production. Nevertheless other indicators could be used.

Using as dependent variable the decision of the peers panel (selected-1; not selected-0) and the bibliometric indicators as independent variables we applied binary logistic regression aimed at determining those indicators that can be used to predict the final decisions made by the peers.

Results

The model

The application of the binary logistic regression lead to the following model:

$$P_i = \frac{e^{-1.88 + 1.116SJR_m + 0.064HCD}}{1 + e^{-1.88 + 1.116SJR_m + 0.064HCD}}$$

where P_i is the probability of the applicant i to be selected by the peers for funding. The SJR_m and the HCD are the indicators that were found to be able to represent the decisions made by the peers panel. The sensitivity determined for this model was 73.2%, the percentage of false positives obtained was 35% and 70% of the cases are predicted correctly by the model. The probability of the forecasted probability by the model for a selected applicant to be higher than that of a non-selected one is 75.3% (ROC curve).

Forecasts

The predictions given by the model are useful in preparing the decisions to be taken by the peers, but the use can be increased if complemented with some type of uncertainty measure. Here, this is shown using the margins concept. Margins are being used in bibliometrics at the individual level for the first time as far as we know.

In Figure 1 is shown the probability of a given applicant to be selected for funding as we increase the value of the HCD and SJR_m , respectively, and maintaining the average value of the other variable. For each predicted value is also shown the confidence interval at 95%, working as the uncertainty measure. All this information can be used by the peers to improve the decision making process.

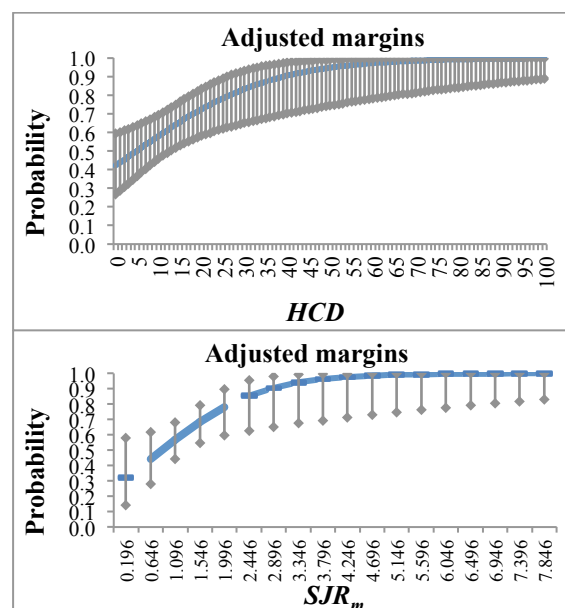


Figure 1. Predicted probabilities complemented with confidence intervals (95%). The dashed zone represents values with a few observations.

Conclusions

From this study some findings can be drawn:

- ✓ The bibliometric indicators are useful in describing the performance of applicants with PhD's earned 6 to 12 years ago.

- ✓ A composite indicator (HCD and SJR_m) when used by the peers will have a positive impact on the final decision.
- ✓ Bibliometric indicators can be used, for example, as input tool helping peers panel in their decision making process as the indicators can give consistent and objective information.
- ✓ The HCD is a serious candidate as tool in support decisions of peer evaluations as it was also found to be useful in describing the final decisions in other types of openings (Vieira et al., 2014a, 2014b).

Acknowledgements

Elizabeth Vieira wishes to acknowledge the financial support from FCT (Foundation of Science and Technology), Portugal, through a grant SFRH/BPD/99246/2013. Data on applicants and results were kindly made available by FCT.

References

- Bornmann, L., & Leydesdorff, L. (2013). The validation of (advanced) bibliometric indicators through peer assessments: A comparative study using data from incites and f1000. *Journal of Informetrics*, 7(2), 286-291.
- Gonzalez-Pereira, B., Guerrero-Bote, V.P., & Moya-Anegon, F. (2010). A new approach to the metric of journals' scientific prestige: The SJR indicator. *Journal of Informetrics*, 4(3), 379-391.
- Moed, H. F. (2005). *Citation Analysis in Research Evaluation*. The Netherlands: Springer
- Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4(3), 265-277.
- Norris, M., & Oppenheim, C. (2003). Citation counts and the research assessment exercise V - archaeology and the 2001 RAE. *Journal of Documentation*, 59(6), 709-730.
- Rinia, E. J., van Leeuwen, T. N., van Vuren, H. G., & van Raan, A. F. J. (1998). Comparative analysis of a set of bibliometric indicators and central peer review criteria - evaluation of condensed matter physics in the Netherlands. *Research Policy*, 27(1), 95-107.
- Vieira, E. S., Cabral, J. A. S., & Gomes, J.A.N.F. (2014a). Definition of a model based on bibliometric indicators for assessing applicants to academic positions. *Journal of the Association for Information Science and Technology*, 65(3), 560-577.
- Vieira, E. S., Cabral, J. A. S., & Gomes, J.A.N.F. (2014b). How good is a model based on bibliometric indicators in predicting the final decisions made by peers? *Journal of Informetrics*, 8(2), 390-405.
- Vieira, E. S., & Gomes, J.A.N.F. (2011). An impact indicator for researchers. *Scientometrics*, 89(2), 607-629.

Differences by Gender and Role in PhD Theses on Sociology in Spain

Lourdes Castelló Cogollos¹; Rafael Aleixandre Benavent², Rafael Castelló Cogollos³

¹ *lourdes.castello@uv.es*

UISYS-Universitat de València. Plaça Cisneros, 4. 46003-València (Spain)

² *rafael.aleixandre@uv.es*

INGENIO (CSIC-Universitat Politècnica de València). UISYS-Universitat de València. Plaça Cisneros, 4. 46003-València (Spain)

³ *rafael.castello@uv.es*

Departament de Sociologia i Antropologia Social. Facultat de Ciències Socials. Universitat de València Av. Tarongers, 4b. 46021 València (Spain)

Introduction

In recent years, there has been a growth in the number of papers that synthesize empirical research studies on gender and sex inequalities in academic statements. Furthermore, these studies can comply with European requirements of equalities since the Treaty of Amsterdam of 1999 enacted that equality between men and women should be included in all policies (Fernández Álvarez, 2014).

Theses are the research papers by excellence and a good indicator to elucidate the lines and research trends in a field of science, since this work must be original and specialized and are subject to a rigorous academic assessment (Delgado López Cózar et al., 2006).

Our objective is to analyse the differences in gender representation in the Spanish sociological theses focusing on three actors involved in the process: PhD students, supervisors and academic assessment boards.

Method

Records were obtained from TESEO, the governmental database of the Spanish Ministry of Education, Culture and Sport, which includes the Spanish theses defended and approved after evaluation. The search was limited to theses indexed by UNESCO codes related to Sociology (code 63) and to theses from the departments of Sociology of Spanish universities. A relational database was created to analyse and compare results.

Results

The total number of theses defended was 3,413. In the role of the PhD student, men presented 253 more theses than women did, while in the role of supervisor and academic assessment board, the differences were much greater: 1,004 and 1,159, respectively (Table 1).

Table 1. Number of PhD theses by gender and role.

Role	Male	Female	Total
PhD student	1,833	1,580	3,413
Supervisor	1,593	589	2,182
Assessment board	1,824	665	2,489

The percentage difference between males and females for PhD students is of 7 points, while for supervisors is of 47 points in favour of males, and for academic assessment boards this difference is of 47 points (Figure 1). The highest percentage of difference occurs in the role of academic assessment board, where 73.3% of board members were of males (Figure 1).

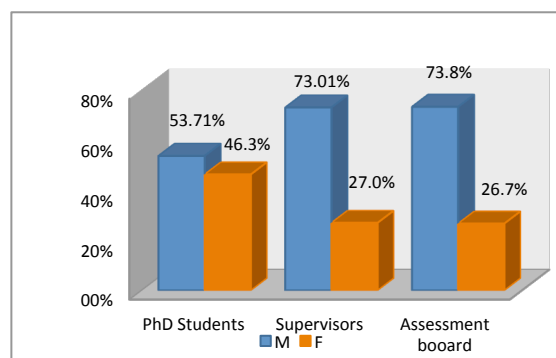


Figure 1. Percentage differences in PhD theses by gender and role

In the annual evolution of the percentages in the roles of supervisor and academic assessment board, men remain between 70% and 80% and women between 20% and 30%. On the contrary, from 2006-2010 period, women-PhD students reach parity (50%) and even surpass men in conducting thesis, ranking 57.8% in the last five-year period analysed (2011-2013) (Figure 2).

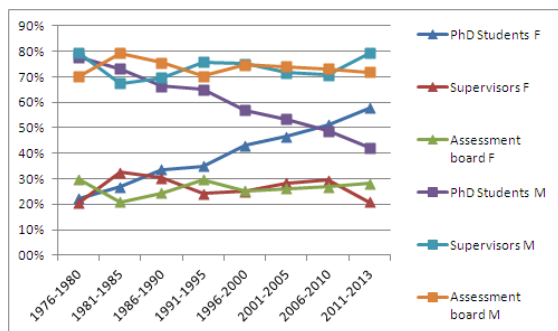


Figure 2. Five-year evolution of PhD theses by gender and role (1976-2013).

Discussion and Conclusions

Although a century has elapsed since the first woman enrolled in a Spanish university and its presence in several strata of the university has greatly improved, the percentage of women compared to men remain far from achieving parity in some roles.

The participation of women at the Spanish universities has increased steadily and its consolidation as PhD students today is a reality (Bermudez et al., 2011). However, from this stage, the academic careers of women slow down and the number of women who leave after doctorate is large (Bordons et al., 2003; Villarroya et al., 2008). Consequently, the percentage of female lecturers in Spain is between 30% and 35%, and the female professors between 14% and 20%. Therefore, it is noteworthy the existing great inequality in the Spanish universities as a professional field and that even though women are more numerous and better prepared than men at all levels of education, this is not reflected in prestigious academic positions (González Alcaide et al., 2009).

In conclusion, the promotion of women to positions of great academic responsibility is slow and is not in line with the number of women who obtained his doctorate in Sociology in Spain. Future research could explore other variables and behaviours, for example, if students of one gender tend to have supervisors from other different gender, as well as these trends in other fields and countries.

Acknowledgments

This work has benefited from assistance by the National R+D+I of the Ministry of Economy and Competitiveness of the Spanish Government (CSO2012-39632-C02-01) and Prometeo Program for excellent research groups of Generalitat Valenciana (GVPROMETEO2013-041).

References

- Bordons, M., Morillo, F., Fernández, M.T., & Gómez, I. (2003). One step further in the production of bibliometric indicators at the micro level: Differences by gender and

professional category of scientists.

Scientometrics, 57(2), 159-173.

- Bermúdez, M.P., Guillén Riquelme, A., Gómez García, A., Quevedo Blasco, R., Sierra, J., & Buela Casal, G. (2011). Análisis del rendimiento del doctorado en función del sexo. *Educación XXI*, 14(1), 17-33.

- Delgado López Cózar, E., Torres Salinas, D., Jiménez Contreras, E., & Ruiz Pérez, R. (2006). Análisis bibliométrico y de redes sociales aplicado a las tesis bibliométricas defendidas en España (1976-2002): Temas, escuelas científicas y redes académicas. *Revista Española de Documentación Científica*, 29(4), 493-524.

- Fernández Álvarez, O. (2014). The gender perspective in managing knowledge through cross-curricular studies in higher education. *Procedia - Social and Behavioral Sciences*, 161(19), 269-274.

- González Alcaide, G., Agulló Calatayud, V., Valderrama Zurián, J.C., & Aleixandre Benavent, R. (2009). Participación de la mujer y redes de coautoría en las revistas españolas de Sociología. *Revista Española de Investigaciones Sociológicas*, 126, 153-166.

- Villarroya, A., Barrios, M., Borrego, A., & Frías, A. (2008). PhD theses in Spain: A gender study covering the years 1990–2004. *Scientometrics*, 77(3), 469–483.

The Trends to Multi-Authorship and International Collaborative in Ecology Papers

João Carlos Nabout¹, Marcos Aurélio de Amorim Gomes², Karine Borges Machado³, Barbbara da Silva Rocha⁴, Meirielle Euripa Pádua de Moura⁵, Raquel Menestrino Ribeiro⁶, Lorraine dos Santos Rocha⁷, José Alexandre Felizola Diniz-Filho⁸ and Ramiro Logares⁹

¹*joao.nabout@ueg.br*, ⁷*lo.rrane@hotmail.com*

State University of Goiás, Br 153, 3105, Fazenda Barreiro do Meio, CP 459, CEP 75132-903, Anápolis, GO (Brazil)

²*marcos.bioamorim@gmail.com*, ⁵*meirielle-euripa@hotmail.com*, ⁶*raquel.menestrino@gmail.com*

State University of Goiás, PPG Recursos Naturais do Cerrado, Br 153, 3105, Fazenda Barreiro do Meio, CP 459, CEP 75132-903, Anápolis, GO (Brazil)

³*karineanjos06@hotmail.com*, ⁴*barbbararocha@hotmail.com*

Federal University of Goiás, PPG Ecologia e Evolução, Campus Samambaia, Goiânia, GO (Brazil)

⁸*diniz@ufg.br*

Federal University of Goiás, Campus Samambaia, Goiânia, GO (Brazil)

⁹*ramiro.logares@gmail.com*

Institute of Marine Sciences, CSIC, Barcelona ES-08003 (Spain)

Introduction

The global number of papers published in different areas has increased over the years (King, 2004). Moreover, the science has experimented changes in academic production scenarios, such as decreased number of solo and increased team authors over the years (Nabout et al., 2015). For many a researcher the number of authors is one measure of collaborations (Price, 1958).

In fact the collaboration has promoted strong changes in science, and there are different reasons for collaboration: increased publication quality (Padial et al., 2010), and sharing costs and ideas (Vermeulen, Parker & Penders, 2013). For Ecology, complex questions such as global climate change, conservation plans of biodiversity among others, have promoted collaboration between scientists (Nabout et al., 2015). Moreover, there are different possible levels of collaboration and an important paper of Katz & Martin (1997) addresses this issue. For these authors, collaboration is: "Thus, a 'research collaboration' could be defined as the working together of researchers to achieve the common goal of producing new scientific knowledge." (Katz & Martin, 1997)

In general, the collaboration can be inter- or intra- at different spatial scales (e.g. national or international; intra or interinstitutional). This variation indicates levels of collaboration. Therefore, collaborations can occur between researchers from the same institution, between institutions of the same country and between different countries (Katz & Martin, 1997). Several

methods have been proposed to measure the collaboration and using different units (researchers, institutes).

The aim of this study is to investigate the temporal trends of number of authors in Ecology journals between 1945 until 2014. Moreover, we will investigate the influence of level of collaboration (intra-institution - II; between-institutions - BI and between-countries - BC) in scientific quality (i.e. number of citation of paper). Our hypothesis is that collaborative papers (BC) generate more citations.

Data

To assess the number of authors and level of collaboration in Ecology papers, we selected all journals listed in category "Ecology" in Web of Science (www.isiknowledge.com, searched in February of 2015). We selected for this study only original articles (type of document), excluding notes, reviews, errata and others. We adopted this strategy to control the influence of type of document in the number of authors (Padial et al., 2010). The selection of papers considered all periods available in the Web of Science database (1945-2014). For collaboration analysis we consider only recent papers (2012-2014). For each paper, the following data were obtained: i) number of authors, ii) number of citations, iii) year of publication, and iv) the level of collaboration. For this last variable, papers were categorized according to the number of institutions of the authors and co-authors and their location. Therefore, authors affiliated with the same

institution were classified as intra-institutional collaboration (II); between-institutional in same country (BI) or between institution in different countries (BC).

Temporal Trends of Number of Authors

We found a total of 333,214 articles published in journals in the Ecology of Thomson-ISI between the years 1945 and 2014. The investigation of the number of authors per paper demonstrated a strong decay in the numbers of single-authored papers. In the early years, about 80% of papers in Ecology were single-authored. In 2014 this value is 4.8%. Statistical models suggest that in 2030 only 0.01% of papers will be single-authored (see Nabout et al., 2015). In addition, the number of papers with two authors have also declined from the beginning of the '90s. Therefore, recently there has been observed the increment in the number of papers with four and five authors, which enhances the tendency of multi-authored papers in Ecology. This trend has been observed in many other areas of science (Abt, 2007).

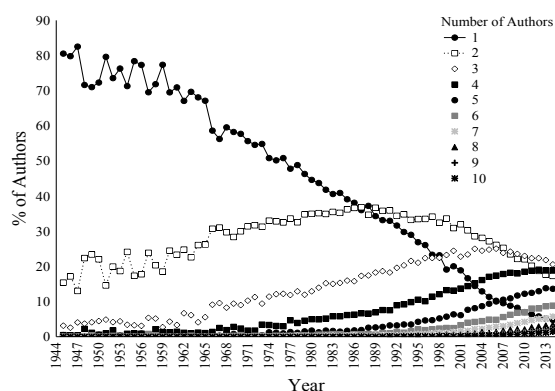


Figure 1. Temporal trends of the proportion of number of authors in Ecology Papers.

Levels of Collaboration

The papers of the years 2011, 2012 and 2013 exclusively of Ecology, totaling 10,457, were classified according to the level of collaboration (II, BI or BC). The Kruskal-Wallis (H) one-way analysis of variance by ranks was performed to assess if the number of citations is affected by the level of collaboration. We found a strong statistically significant difference ($P < 0.01$), suggesting that collaborative papers written by authors from different countries received more citations Figure 2). This result reinforces the importance (and a recent trend) of international collaboration.

Using the same analysis we observed that the number of authors differs significantly between the levels of collaboration. In other words, BC papers have higher number of authors than those of SI and BI papers ($H = 1868$, $P < 0.001$). Therefore, the

number of authors can also be an indication of the level of collaboration.

Finally, our work shows an increase in the number of multi-authored papers in Ecology. This is probably due to the complexity of questions in ecology which promotes collaboration between researchers. In addition, international collaborations have promoted papers with more citations (see Glänzel, 2001). Thus, the reduction of travel costs and the internet has allowed greater exchange between countries. In addition, governmental strategies can help in the exchange of researchers, such as the Program Science Without Border in Brazil. Thus, we encourage collaboration between researchers seeking to improve the ecological research of countries.

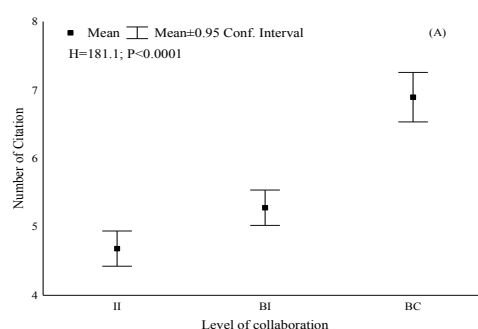


Figure 2. Number of citations for each one of level of collaboration.

Acknowledgments

Our work on Scientometrics and Ecology has been continuously supported by different grants FAPEG, CNPq and CAPES.

References

- Abt, H.A. (2007). The future of single-authored papers. *Scientometrics*, 73(3), 353–358.
- Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics*, 51(1), 69–115.
- Katz, S., & Martin, B.R. (1997). What is research collaboration? *Research Policy*, 26(1): 1–18.
- King, D.A. (2004). The scientific impact of nations. *Nature*, 430, 311–316.
- Nabout, J.C., Parreira, M.R., Teresa, F.B., Carneiro, F.M., Cunha, H.F., Onde, L.S., Caramori, S.S. & Soares, T.N. (2015). Publish (in a group) or perish (alone): the trend from single- to multi-authorship in biological papers. *Scientometrics*, 102, 357–364.
- Padial A.A., Nabout, J.C., Siqueira T., Bini, L.M., Diniz-Filho, J.A.F. (2010). Weak evidence for determinants of citation frequency in ecological articles. *Scientometrics*, 85, 1–12.
- Price, D.J.S. (1963). *Little Science, Big Science*. New York: Columbia University Press.
- Vermeulen, N., Parker, J. N., & Penders, B. (2013). Understanding life together: A brief history of collaboration in biology. *Endeavour*, 37(3), 162–171.

A Bootstrapping Method to Assess Software Impact in Full-text Papers

Erjia Yan¹ and Xuelian Pan²

¹*ey86@drexel.edu*

Drexel University, College of Computing and Informatics, 3141 Chestnut Street, Philadelphia, PA 19104
(U.S.A.)

²*panxuelianmail@gmail.com*

Nanjing University, School of Information Management, Nanjing, 210093 (P.R. China)

Introduction and Motivation

There is a concerted effort to study science of science in multiple spheres. However, a clear gap exists in how to incorporate digital outputs, such as software, as an integral component in scholarly communication. This tension has become aggravated in recent years because software can be the end products in many scientific inquiries. Therefore, there is the need to build a framework to assess the impact of software in science. One cornerstone in the framework is the design of text-based methods to identify software entities in full-text corpora because these entities are largely mentioned in the text rather than formally cited in the way as their publications counterpart. This research-in-progress paper will serve this purpose by the development and evaluation of a bootstrapping method to automatically extract software entities from a full-text data set.

Despite the effort of indexing digital outputs such as Thomson Reuters' Data Citation Index or SageCite by University of Bath, U.K., the use of full-text data is necessary to identify patterns of software references because these digital outputs are referenced in unsystematical ways in scientific literature. They can be embedded in documents by digital object identifiers (DOIs), hyperlinks, and featured on dedicated websites or simply be mentioned in paragraphs, footnotes, endnotes, acknowledgements, or supplementary materials. A 2014 citation study on three oceanographic data sets showed that these digital outputs are more likely to be mentioned in the text than formally cited (Belter, 2014). Intuitively, one would think of curating a list of software names; however, it will not be feasible due to the velocity, variety, and volume of software that has been developed and applied constantly. Thus, merely using metadata or static listings is incapable of capturing the full extent of the impact of software. Instead, full-text publication data provide the crucial context for this purpose.

This study will use a bootstrapping method to identify software uses in a full-text data set. It will allow us to expand the impact and attribution mechanism by assessing the impact of software.

Methods

The bootstrapping method is used to extract software entities from full-text papers. It is a self-sustaining technique used to iteratively improve a classifier's performance through seed terms (Riloff & Jones, 1999; Riloff, Wiebe, & Wilson, 2003). The bootstrapping process contains the following steps: (1) Label seed terms or learned entities in the text. Seed terms are used in the first iteration, and learned entities are used in other iterations. (2) Generate contextual patterns of seed terms in the first iteration, and create contextual patterns of learned entities in other iterations. (3) Score these contextual patterns and select top ranked N patterns as candidate patterns. (4) Score entities extracted by candidate patterns and select top ranked M entities as learned entities. (5) Go back to the first step until the system cannot learn any new positive entities.

The calculation of pattern scores and entity scores determine the effectiveness of the bootstrapping method. If a pattern gets a higher score, then it is selected into the candidate pattern pool. Entities extracted by these candidate patterns are considered as candidate entities. To boost the performance, we incorporated three heuristic rules to the calculation of pattern scores. The first feature is an unlabeled entity containing at least one uppercase letter. An entity with this feature gets a score of 1 if it contains one or more uppercase alphabetic letters; otherwise, it gets a score less than 1. The second feature focuses on version numbers. An entity with this feature gets a score of 1 if a version number is collocated. The third and fourth features deal with the presence of trigger words: a score of 1 if the left context (third feature) or right context (fourth feature) of an entity contains trigger words.

Preliminary Results

To construct a corpus that has a good balance between sentences having software entity that mentions and does not mention, we selected 427 sentences that a particular software entity is mentioned from papers published between January 6 and December 29, 2013 in the data set. 573 sentences that do not contain software entities were also included in the corpus. We use this data collection method to attain a balanced experiment set to evaluate several entity extraction methods.

Experiments that use randomly sampled sentences will be pursued as future work. We used nine frequently occurring seed terms in the proposed bootstrapping method, including SAS, SPSS, MotIV, PAML, rGADEM, Limma, PICS, PHYLIP, and Minitab. To prepare the gold standard, we manually labeled software entities in the experiment data set and in total annotated 292 unique entities. The annotations are considered as the gold standard.

Table 1 displays the experimental results of the RlogF metric entity extraction system (Thelen & Riloff, 2002), Stanford Pattern-based Information Extraction and Diagnostics (SPIED), and our software extraction system. All methods in Table 1 used the same sets of seed terms, stop word list, and common word list.

Table 1. Experimental results of software extraction.

<i>System</i>	<i>Prec</i>	<i>Recall</i>	<i>F</i>
RlogF	91%	7%	0.12
SPIED	40%	28%	0.33
OurSystem	80%	62%	0.70

Table 1 shows that our system performed better than RlogF and SPIED based on the F score. Although RlogF has the highest precision, it missed a great number of software entities and resulted in the lowest recall. By comparing the software entities extracted by our system and the gold standard, we found seven of the one-time occurring entities were not identified by our system thus reducing the recall. We speculate that the recall may be improved when more sentences that contain low frequently occurring software entities are added to the data set such that the bootstrapping method will be able to learn their contexts.

Table 2. Popular software use in science.

<i>Freq</i>	<i>Software entities</i>
2	Prism, PASW, Vienna RNAfold, survival, Stata, SeqMan, rtracklayer, R2WinBUGS, Quantity One, PyPop, Origin, Microsoft Office Excel, JMP, GeneSpring GX, genefilter, FlowJo, Effective T3, Cytoscape, COMSTAT, CellquestPro, APE, ADE4, MetaMorph Imaging System
3	SigmaPlot, WinBUGS, T3SEpre, Statistica, MetaMorph, TiMAT2, stats, Statistical Package for the Social Sciences, STADEN, limma Bioconductor
4	HyPhy, IRanges, ImageJ, Affy, Vienna RNA
5	SigmaStat, MEGA, Vegan, Geneious
≥6	R, SAS, SPSS, MotIV, Bioconductor, Weka, PAML, rGADEM, Limma, PICS, PHYLIP, Minitab, Cellquest, RNAfold, Image J, GraphPad Prism

Table 2 shows 59 popular software entities in science which occurred more than once in the test corpus based on our extraction method. Statistical software packages are well presented in Table 2; however, we also see some domain-specific open access software tools—future impact assessment may primarily focus on these.

Conclusion and Future Work

The contemporary research landscape is changing: software has increasingly been developed and applied in many data-driven projects. Therefore, there is the need to assess its impact on science and to incorporate software in scientific evaluations. This paper is part of a larger effort to build a scientific assessment framework for digital outputs that include software and data. It has proposed a bootstrapping method to extract software entities in a full-text corpus. Results show that it has successfully extracted software entities with the F score at the 0.7 level which is an improvement over the baseline methods RlogF and SPIED. Future work will involve using the whole *PLOS ONE* full-text set and introducing more advanced features to further enhance the performance of the method. Research will also benefit from integrating the number of full-text software entity mentions with citation- and usage-based metrics to complement the impact assessment of software.

Acknowledgments

Erjia Yan is supported by the National Consortium for Data Science (NCDS) Data Fellows program for the project “*Assessing the Impact of Data and Software on Science Using Hybrid Metrics*”. Xuelian Pan was a visiting PhD candidate at Drexel University, supported by China Scholarship Council, when this work was performed.

References

- Belter, C. W. (2014). Measuring the Value of Research Data: A Citation Analysis of Oceanographic Data Sets. *PLOS ONE*, 9(3), e92590.
- Riloff, E., & Jones, R. (1999). Learning Dictionaries for Information Extraction by Multi-Level Bootstrapping. *Proceedings of AAAI-99*. Menlo Park, CA: The AAAI Press.
- Riloff, E., Wiebe, J., & Wilson, T. (2003). Learning subjective nouns using extraction pattern bootstrapping. *HLT-NAACL Association for Computational Linguistics*, (pp. 25-32).
- Thelen, M., & Riloff, E. (2002). A bootstrapping method for learning semantic lexicons using extraction pattern contexts. *Proceedings of the ACL-02 Association for Computational Linguistics*, (pp. 214-221).

Article and Journal-Level Metrics in Massive Research Evaluation Exercises: The Italian Case

Marco Malgarini¹, Carmela Anna Nappi¹ and Roberto Torrini¹

{marco.malgarini, carmelaanna.nappi, roberto.torrini}@anvur.it
ANVUR, Via Ippolito Nievo 35, 00153, Rome. (Italy)

Introduction

Article level metrics are usually the preferred choice for research evaluation. However, for recent articles they may be integrated or substituted considering some measure of journal impact (Abramo et al., 2012). The use of journal level metrics is also often considered as particularly appealing for administrative purposes, because of their readily availability, easiness to use and comprehensibility (Bordons et al., 2002). On the other hand, the IF is often criticized on the grounds of its possible biases and lack of methodological consistency (Vanclay, 2012). The aim of our paper is to provide evidence about the effects of the use of journal level metrics on the results of a massive research evaluation exercise like the one that has been performed in Italy with reference to the period 2004-2010 (VQR 2004-2010, see Ancaiani et al., 2015). More specifically, in the following we evaluate the effects of the use of the impact factor (IF) on the ranking of Italian Universities at the aggregate level, at the area level and for individual researchers.

Effect of the use of the Impact Factor at the University level

In order to assess the impact of the use of IF, we calculate two different indicators of research quality, denoted as R_VQR and R_IF . The former is based on the rules used for the VQR, and the latter uses only the Impact Factor in order to evaluate the articles; the analysis is limited to the research products evaluated only with bibliometrics. We then rank the 93 Italian Universities on the basis of those indicators, finding that the Spearman correlation index among the two rankings is equal to 0.92; moreover, the R^2 of a regression of R_VQR over R_IF and a constant is equal to 0.85. Hence, the analysis at the aggregate level shows that the final ranking of Italian Universities based on journal metric alone is very close to that obtained with the VQR algorithm (see also Figure 1).

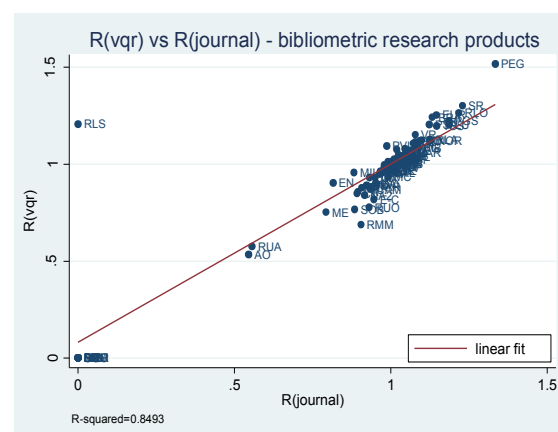


Figure 1 – The relationship among University evaluation performed with different metric.

Effect of the use of the Impact Factor at the Area level

However, it is well possible that the relationship is weaker when we are interested in ranking Universities in each scientific area. In order to shed light on this issue, we repeat the analysis for the 14 areas considered in the VQR (Table 1). Correlation between the two rankings is still above 0.8 in all the Research Areas except for Chemistry. The Spearman correlations among rankings are significant at 5% level in all the research areas.

Table 2 reports the coefficients of the regressions of R_VQR on R_IF (beta) and a constant (alpha); the table also reports the R^2 of the regression (column 3) and the standard deviation (column 4) normalized with respect to the average value of R_{iVQR} in each Area. Standard deviation is pretty low if compared to the average value of R (around 7%) in the Areas of Mathematics, Physics and Industrial Engineering, while in Earth Science, Medicine and Biology the normalized standard deviations grow to 17% of the average level of R in those areas. Similarly, the areas with a low normalized standard deviation are also those with a higher R^2 and vice-versa. Hence, results confirm that the two evaluation methods bring very similar results also at the area level.

Table 1. Spearman Correlation between Rankings obtained with VQR bibliometric rules and Journal metric (* indicates statistical significance at 5%).

Research Area	Spearman	# Univ.
Mathematics	0.926*	64
Physics	0.825*	65
Chemistry	0.654*	60
Earth Science	0.724*	46
Biology	0.861*	66
Medicine	0.701*	58
Veterinary Sciences	0.876*	50
Construction engineering	0.720*	54
Industrial engineering	0.769*	67
Psychology	0.764*	61

Table 2. Sensitivity of research evaluation to the use of the Journal Impact Factor at the area level.

	(1)	(2)	(3)	(4)
Research Area	α	β	R^2	St. dv.
Mathematics	-0.055	1.039***	0.921	0.058
Physics	-0.13**	1.124***	0.847	0.060
Chemistry	-0.029	0.998***	0.706	0.100
Earth Science	0.180	0.815***	0.478	0.170
Biology	-0.142	1.132***	0.720	0.168
Medicine	0.083	0.894***	0.340	0.167
Veterinary Sciences	-0.004	1.016***	0.787	0.125
Construction engineering	0.186*	0.813***	0.532	0.100
Industrial engineering	-0.014	1.004***	0.675	0.070
Psychology	0.0778	0.916***	0.744	0.155

Effect of the use of the Impact Factor at the individual level

We finally look at how the use of the IF influences evaluation results for each individual researcher. In this case, we regress individual scores obtained using either citations or the Impact Factor. Results of the estimation are reported in Table 3.

The relationship among the results obtained with the two different metrics is now rather weak: the R^2 of the regression is equal to 0.18 for the whole sample, varying between 0.20 and 1.156 in each year. The constant of the regression is rather high, while the beta coefficient associated with the IF is much lower than in previous estimates. Hence, at the individual level using alternatively only the citations or only the impact factor would imply a rather different outcome.

Table 3. Citations vs Journal Metric scores at individual level.

	Coefficient							
	Whole sample	2004	2005	2006	2007	2008	2009	2010
IF	0.488 ***	0.525 ***	0.531 ***	0.521 ***	0.507 ***	0.487 ***	0.507 ***	0.383 ***
Constant	0.280 ***	0.247 ***	0.232 ***	0.254 ***	0.282 ***	0.301 ***	0.233 ***	0.374 ***
#obs.	76,15	9,23	9,77	10,24	10,88	11,56	12,15	12,31
R^2	0.184	0.201	0.197	0.202	0.197	0.194	0.186	0.156

Conclusions

Overall, results may be considered as supportive of the idea of using two different bibliometric indicators for assessing research quality: on one hand, the use of the IF is not found to bias in a significant way University rankings, both at the aggregate and at the Area level; on the other hand, at the individual level, citations and IF evaluation are found to be rather different, pointing to the need of integrating the two different information in order to obtain a more robust measure of research quality for each individual researcher.

References

- Abramo G., D'Angelo C.A., & Costa F. (2012). Citations versus journal impact factor as proxy of quality: could the latter ever be preferable? *Scientometrics*, 84(3), 821-833.
- Ancaiani A. et al. (2015). Evaluating scientific research in Italy: the 2004-2010 Research evaluation exercise. Forthcoming in *Research Evaluation*.
- Bordons, M., Fernández, M. T., & Gomez, I. (2002). Advantages and limitations in the use of impact factor measures for the assessment of research performance. *Scientometrics*, 53(2), 195-206.
- Vanclay J.K. (2012). Impact factor: outdated artefact or stepping-stone to journal certification? *Scientometrics*, 92(2), 211-238.

Accounting for Compositional Effects in Measuring Inter-Country Research Productivity Differences: The Case of Economics Departments in Four European Countries

Giannis Karagiannis¹ and Stelios Katranidis²

¹ karagian@uom.edu.gr, ² katranid@uom.edu.gr

University of Macedonia, Department of Economics, Egnatia str. 156, 546 36Thessaloniki (Greece)

Introduction

Most of cross country studies on research productivity differences do not take into account compositional differences in academic staff force, such as sex, years of experience, origin of PhD studies, even though there are well documented evidence that (a) males tend to publish more than females (Gupta et al., 1999); (b) junior academic staff tend to publish more and in better outlets than senior stuff (Ben-David, 2010); and (c) academic staff with PhD studies in North America tend to be more productive (Katranidis et al., 2014). These aspects of observed faculty heterogeneity affect research productivity and are expected to have an impact on country average performance (Combes et al., 2003)¹.

Methodology and Data

In this paper we use the pure output or the single constant input DEA model, which is also known in the literature as the Benefit-of-the-doubt (BoD) model, to construct in the first stage a composite indicator of research productivity based on publication and citation counts at the faculty staff level. In particular, the BoD model in its multiplier form is given as (Cherchye et al., 2007):

$$I^k = \max_{s_i^k} \sum_{i=1}^N s_i^k I_i^k$$

$$\text{st } \sum_{i=1}^N s_i^k I_i^j \leq 1^j \quad \forall j = 1, \dots, K \quad (1)$$

$$s_i^k \geq 0 \quad \forall i = 1, \dots, N$$

where I_i^k is the i^{th} sub-indicator of the k^{th} unit, s_i^k are the weights to be estimated, j is used to index units and i to index sub-indicators which in our case correspond to different research outcomes (i.e., publication and citation counts). The BoD model is equivalent to the multiplier form of the input-oriented, constant returns to scale (CRS) DEA model when there is a single constant input that takes the value of one for all evaluated units. Based

on this, the dual formulation of the BoD model is given as:

$$I^k = \min_{\lambda_j^k} \sum_{j=1}^K \lambda_j^k 1^j$$

$$\text{st } \sum_{j=1}^K \lambda_j^k I_i^j \geq I_i^k \quad \forall i = 1, \dots, N \quad (2)$$

$$\lambda_j^k \geq 0 \quad \forall j = 1, \dots, K$$

where λ refers to intensity variables. Then the results at the country level are obtained by using the aggregation rule suggested by Karagiannis (2013), namely:

$$I = \frac{1}{K} \sum_{k=1}^K I^k \quad (3)$$

Thus, the aggregate composite performance indicator equals the simple (un-weighted) arithmetic average of the estimated individual composite indicators.

At the second stage we use Ray (1991) regression model to account for several contextual variables such as country dummies, a sex dummy, years of experience, and origin of PhD studies (i.e., overseas, Europe, home country and inbreeding), i.e.:

$$I^k = h(z_r^k) + e^k, \quad (4)$$

where r is used to index contextual variables and is $e^k < 0$ represents managerial inefficiency pure of (favorable and unfavorable) contextual variables. After taking into account the impact of contextual variables through (4) we re-calculate faculty level research performance scores and country averages. Our interest is to examine if and by how much these country averages differ from the unadjusted ones obtained via (1) or (2), and which countries are affected the most by the contextual variables.

We apply the above methodology to European faculty members in selected departments of Economics. In particular our sample consists of four countries, i.e., Belgium, Denmark, Greece and Portugal and a total of 383 faculty members and 15 departments. The analysis covers the period 1996-

¹ This research is implemented through the Operational Program "Education and Lifelong Learning" and is co-financed by the EU (European Social Fund) and Greek national funds.

2012 and the publication and citation count data come from Scopus database.

Empirical Results

Our main empirical results are summarized in the following tables:

Table 1. Unadjusted Composite indicator vs. efficient and unproductive faculty members.

	Unadjusted Composite indicator	Number of efficient faculty members	Number of unproductive faculty members
Belgium	0.144	1	6
Denmark	0.105	0	10
Greece	0.084	0	9
Portugal	0.062	1	18

Table 2. Number of unproductive faculty members vs. Adjusted Composite Indicator.

	Number of unproductive faculty members	Max value	Standard deviation	Adjusted Composite Indicator
Belgium	6	1	0.18	0.120
Denmark	10	0.588	0.11	0.100
Greece	9	0.667	0.10	0.086
Portugal	18	1	0.13	0.062

According to the unadjusted composite indicator, Belgian faculty members are found to be the more efficient and Portuguese the less efficient. In addition, in these two countries we can find the two fully efficient faculty members we have identified. At the same time these two countries are the ones with the relatively higher heterogeneity in terms of research productivity as indicated by the standard deviation of the unadjusted composite indicator.

When the composite indicator scores are adjusted for the potential impact of the aforementioned contextual variables by means of (4), the resulting efficiency scores change but not as much. They tend to improve a little bit for Belgium, Denmark and Portugal because these countries have a relatively higher percentage of inbred faculty members who in turn perform better compared to other faculty members. On the other hand, Portugal performance is adversely affected by the relatively larger percentage of females (31%) who though publish less than males and this counteract with the positive effect of inbred faculty, resulting in an unchanged national average.

Concluding Remarks

The empirical results indicate that the overall effect of the contextual variables considered is positive for the two northern European countries, i.e. Belgium and Denmark, and negligible for the two

southern European countries, i.e., Greece and Portugal. Nevertheless, the two northern European countries perform better than the two southern European countries, regardless of environmental differences.

References

- Ben-David, D. (2010). Ranking Israel's economists, *Scientometrics*, 82, 351-364.
- Cherchye, L., Moesen, W., Rogge, N. & T. van Puyenbroeck. (2007). An Introduction to "Benefit of the Doubt" Composite Indicators, *Social Indicators Research*, 82, 111-45.
- Combes, P.P. & Linnemer, L. (2003). Where are the economists who publish? Publication concentration and rankings in Europe based on cumulative publications. *Journal of the European Economic Association*, 1, 1250-1308.
- Gupta, B.M., Kumar, S., & Aggarwal, B.S. (1999). A comparison of productivity of male and female scientists of CSIR, *Scientometrics*, 45, 269-289.
- Katranidis, S., Panagiotidis, T., & Zontanos, C. (2014). An evaluation of the Greek universities' economics departments. *Bulletin of Economic Research*, 66(2), 173-182.
- Karagiannis, G. (2013). On Aggregate Composite Indicators, unpublished manuscript.

Metrics 2.0 for Science 2.0

Isidro F. Aguillo

isidro.aguillo@csic.es

Cybermetrics Lab. IPP-CSIC. Albasanz, 26-28. Despacho 3E14. Madrid 28037 (Spain)

Science 2.0

The concept Science 2.0 is a recent development designed to take advantage of the new sharing technologies and social networks of the Web 2.0 and that it is now strongly linked to the current and future research policies of the European Commission.

According to ideas developed by Ben Shneiderman this Science in Transition can be described according to two groups of actions,

Integrating the whole research cycle and its stakeholders, including all and both activities and people involved in them, far beyond that focusing only on the authors of papers, and

Opening the whole set of data; tools, results and metrics derived from the cited research (and communication) cycle from the very first moment the information is generated.

The urgent need to adapt the current set of quantitative indicators to this new concept is the reason for this poster. We intend to provide a critical analysis of the current status of the bibliometrics y related quantitative techniques for science evaluation and to introduce a new umbrella term, Metrics 2.0, for describing future scenarios for the discipline.

Current Metrics situation

A SWOT analysis is introduced for describing major issues related to bibliometrics and the attitude of bibliometricians and the rest of scientists' attitude regarding the discipline.

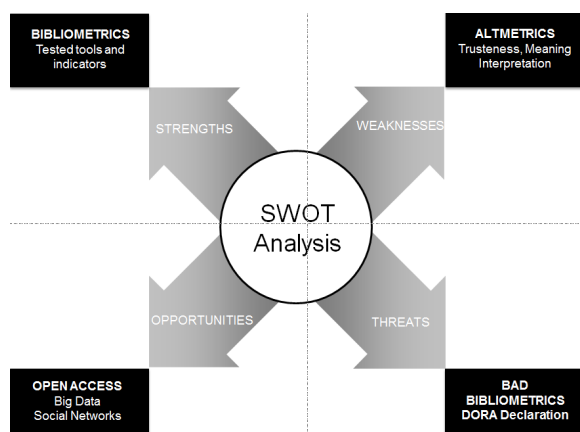


Figure 1. SWOT analysis of bibliometrics.

In recent years the term Informetrics become popular for describing an extended set of disciplines that are closely related to bibliometrics, including patentometrics or webometrics. However the fast development of the Internet, especially regarding the social networks, make this term become obsolete for describing a increasingly complex situation.

Specifically, there are two current developments that are having an impact on the discipline:

Altmetrics 'revolution'. The Web 2.0 tools have used as sources for extracting quantitative data when they are proxies for scholarly communication. Thousands of papers are exploring the capabilities of the different social networks using citation analysis for comparative purposes with mentions, readings or visits to bibliographic units.

Moving beyond Journal-level Metrics. After decades of criticism, and with the recent publication of the Declaration of San Francisco (DORA) the level of analysis is moving from Journal-level to Article-level metrics.

Proposals for Metrics 2.0

Regarding bibliometrics

The most serious problem is related to the way the contribution of each author (and the institution/s to whom is affiliated) is measured in a co-authored document. Traditionally two options were used: Full count (100% of merit for each author) and fractional counting (dividing full merit by the number of authors equally). As the number of authors per paper is growing exponentially, the last option is being discarded in most of the cases. Other alternatives, like identifying in the signature the relative contribution of each author, are still not a feasible option.

Traditionally full count is supported as it favours collaboration, especially international one. But this option is masking relevant phenomena for policy decisions. For example asymmetric collaboration with developing countries provides to their scientists and institutions with output/impact values that are not correlated with their low R&D investment prompting funders to not increase their budgets. Even with symmetric collaboration the full count based results are not able to discover the impact of the current economic crises that reduced

considerably the money invested in scientific research.

Taking into account that is a temporary proposal that intends not to reduce the level of scientific collaboration we suggest using a variant of the full count giving 50% of the merit instead of 100% to each author in papers with two or more authors.

In the case of organizations (and countries) where it is possible to identify the leading institutional author this should be granted the 100% authorship.

Although not a perfect or definitive solution this proposal should be especially useful for solving the problem of 'bad bibliometrics' that spoiled the major university rankings.

Regarding altmetrics

Apart of an ugly name, altmetrics is a confusing tangled set of mixed value tools. A first proposal could be to segregate the field in different subfields according to the tool that is involved. So, twittermetrics is different in both methodology and results interpretation to wikimetrics, for example. But there are two actions that are perhaps far more justified. It is highly recommended to set up a new discipline called Usagemetrics for the analysis of visits, visitors and their behaviour to academic and scientific websites. This is a very rich environment with dozens of candidate variables to build indicators independent from the standard citation motivations. The second moving is related to the tools where mention motivations are close to the citation ones, the most obvious one is Mendeley. In similar cases the proposal is to transfer these tools from altmetrics to the traditional bibliometrics arena.

Regarding Open Data y Big Data

The scientific community is strongly pushing for making openly available the data obtained from the experiments that is used later for preparing papers. Beyond the usefulness of this Open Data for replicating the results or for comparative purposes, the success of the initiative can make available huge amounts of information that could be considered, regarding the size-related challenges they pose, at the same level of the Big Data produced by the so-called Big Science. This is call for the scientific authorities for considering offering Big Data facilities and services for an extended group of scientists.

$$\text{Big Data} = \sum \text{Open Data}$$

Regarding Author Profiles

Until very recently the author-level metrics were technically a complex work when huge numbers of researchers were involved. Now the profiling services offered by several services (ResearcherID,

Google Scholar Citations) or the major interests by the own research organizations (CRIS) and supported by disambiguation identifiers (ORCID) are changing completely the situation. In this new scenario, inspired by the results of the EU Project ACUMEN, we propose to set up author profiles with the following characteristics:

Bibliometric indicators from several sources, Non-bibliometric indicators, like those from altmetrics sources; context information like academic age, academic status, gender, levels of funding, networks membership and role, geographical or discipline biases, among others.

Rankings are a valuable tool if context is appropriately included in their elaboration. Relative indicators (percentages, quartiles) are being shown as far more trusted for this kind of classifications. However the use of composite indicators is still an open unresolved question that is still strongly criticised by the experts.

Conclusion

Metrics 2.0 should open and transparent, with data and indicators provided in a rich metadata environment.

Multiple sources and indicators are required, reflecting the diversity of the research activities, counting correctly and exhaustively the results and evaluating the different levels and magnitudes of the visibility and impacts of these results for all the communities, academic or not.

Presentation of the indicators, including friendly visualization of data is also relevant, but it is probably secondary to offer to end-users unrestricted customisation (including exporting in several formats) capabilities.

Summarising, bibliometricians can no longer been accountants able to extract, standardize, group and visualize the records from the Web of Science, but experts in several fields, with strong knowledge of different information sources and professionals capable of understanding specific needs and contexts ready to customise procedures according to the specific situation. Data, methodology, results and reports should be open to third parties in a mandatory way.

Evolution of Research Assessment in Lithuania 2005–2015

Saulius Maskeliūnas¹, Ulf Sandström² and Eleonora Dagienė³

saulius.maskeliunas@mii.vu.lt, ulf.sandstrom@indek.kth.se, eleonora.dagiene@vgtu.lt

¹Vilnius University Institute of Mathematics and Informatics, Akademijos g. 4, LT-08663 Vilnius (Lithuania)

²KTH, Indek – Department of Industrial Economics and Management,
Lindstedtsvägen 30, 10044 Stockholm (Sweden)

³Vilnius Gediminas Technical University, Sauletekio al. 11, LT-10223 Vilnius (Lithuania)

Introduction

Traditionally, governmental funding of scientific research has been based on input factors (e.g. student numbers), however since the end of the 1980s most developed countries have introduced assessment systems based on scientific output. Numerous examples of research quality assessment can be named as products of innovation and incremental change (Barker, 2007; Hicks, 2012; RDI Council, 2013). An overview of assessment methods applied in Eastern European countries in the field of Social Sciences and Humanities has recently been presented (Pajic, 2015), but information about Lithuanian assessment of research is lacking. Here, we analyse seven sequential Lithuanian methods of research assessment in the period 2005–2015, their influence and consequences.

Evolution of Lithuanian research assessment methodologies

The methodologies of research assessment in Lithuania have changed very often over the period 2005–2015. There is quite a great difference between assessment of papers in Social Sciences & Humanities (SSH) and papers in Science & Technology (S&T). While SSH researchers should have publications in any peer-reviewed journals (Table 1), S&T papers have higher requirements: to gain scores, they have to be published in journals included in *Thomson Reuters Web of Science Core Collection* (WoS) (Table 2).

The value of each research article published in a journal indexed by WoS in SSH was calculated by the following formula in 2006 only:

$$AIV = PVV \frac{N_{IA}}{N_A} \left(1 + \frac{IF_j}{IF_{AIF}} \right) \quad (1)$$

here: *AIV* – contribution of institution authors; *PVV* – [primary] value of unit in points; *N_{IA}* – number of authors from the institution; *N_A* – total number of authors; *IF_j* – journal Impact Factor (Thomson Reuters Journal Citation Reports), *IF_{AIF}* – Aggregated Impact Factor of the subject category in which this journal is listed or average of Aggregated Impact Factors of all subject categories in case the journal is listed in more than one category in Thomson Reuters Journal Citation Reports.

The value of each research article published in a journal indexed by WoS in S&T (2003–2015) and SSH (2008 and 2015) is calculated by the similar formula:

$$AIV = PVV \frac{N_{IA} \sqrt{N_{IP}}}{N_A} \left(1 + k \frac{IF_j}{IF_{AIF}} \right) \quad (2)$$

here: *N_{IP}* – number of different foreign affiliations (but, if *N_{IP}* > *N_A*, then there is considered that *N_{IP}* = *N_A*); *k* = 1 for evaluation until 2007, and *k* = 2 for evaluation of 2008 and later years;

Significant and frequent changes in the evaluation criteria were caused by the search for most fair distribution of governmental funding for Lithuanian research by the Ministry of Science and Education, in order to encourage the highest-level academic research.

All systems of research assessment since 2006 have encouraged S&T researchers to publish their papers in high impact journals and have urged Lithuanian journals to improve their quality as well as actively seek to be indexed in international databases and especially in Thomson Reuters Web of Science. When Thomson Reuters started the expansion of the Web of Science in 2007–2009, many Lithuanian (LT) journals were included into its databases. But, the methodologies used in 2010 and 2011 were disadvantageous to most LT journals as they didn't fulfil the requirements asking *only* for papers in journals which had more than 20 % of citations from journals (citing side) with an impact factor (IF) higher than the aggregate impact factor (AIF) of the respective subject field. This requirement was probably not field neutral but, instead it seemed to be disadvantageous to some fields of science and created funding for other fields. Consequently, some subject fields were downgraded by this requirement and received no score or low scores. However, this citation requirement was not used for evaluation starting from 2012 and will formally withdrawn in 2015.

Since 2009 for SSH and from 2010 for S&T, expert evaluations (by national experts) of papers and monographs presented by institutions is used in addition to previous bibliometric evaluation. Since 2010 the number of 1st level papers and monographs presented by academic and research institutions for expert evaluation is proportional to number of full time equivalent of PhD researchers in both S&T and SSH (i.e., it could be presented not more than one 1st level publication per 5 full time researchers in a research area, and if the unit has doctoral studies in a research area – it can present 1st level publication not depending on number of researchers).

From 2011 the assessment system is carried out every third year (not annually as before). That helps aca-

demographic and research institutions to minimize the drawbacks of productivity fluctuations. The last assessment period was 2009–2011. In 2015, there will be an evaluation of 2012–2014; which will determine the allocation of budgets for 2016–2018 for all universities and governmental research institutions. However it is rather complicated to evaluate the dynamics because of rather frequent changes in evaluation criteria. The benchmarking of Lithuanian research 2009–2013 was run on April 2014 – April 2015 by the Research and Higher Education Monitoring and Analysis Centre (MOSTA), following the methodology prepared by Technopolis Group and involving only international European experts. Here the experts have noticed the need for greater internationalization of Lithuanian Social Science research.

Conclusions

The shift in methodologies for formal assessment of scientific publications produced by Lithuanian higher education and research institutions has urged researchers to communicate their results in international scientific journals, and for the Lithuanian scientific journals to seek inclusion in international databases (especially Thomson Reuters Web of Science, Journal Citation Reports) and to improve their quality. The effect of changes in journals' indicators up until 2012

is the focus of a parallel poster presentation (Dagiene & Sandström, 2015). Whether the introduction of national expert evaluation will change this overall pattern or not is yet to be investigated.

References

- Barker, K. (2007). The UK Research Assessment Exercise: the evolution of a national research evaluation system. *Research Evaluation*, 16(1), 3–12.
- Dagiene, E. & Sandström, U. (2015). Dynamics between National Assessment Policy and Domestic Academic Journals. Poster presentation submitted to *ISSI 2015*.
- Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, 41(2), 251–261.
- RDI Council (2013). Methodology of Evaluation of Research Organizations and Evaluation of Finished Programmes (valid for years 2013–2015). Retrieved on March 20, 2015 from <http://www.vyzkum.cz/FrontClanek.aspx?idsekce=695512>
- Pajic, D. (2015). Globalization of the social sciences in Eastern Europe: genuine breakthrough or a slippery slope of the research evaluation practice? *Scientometrics*, 102(3), 2131–2150.

Table 1. Shift in criteria used for Lithuanian research papers assessment in Social Sciences and Humanities.

2005		2006		2008		Assessment categories	2009		2010; 2011		2015	
Requirements	Value, points	Requirements	Value, points	Requirements	Value, points		Requirements	Value	Requirements	Value	Requirements	Value
Papers in internationally recognised journals	20 (24*)	Papers in publications indexed by Thomson Web of Science Papers in internationally recognised journals Papers in other peer-reviewed journals	30 (S)* 20 (H)* 10 5	Thomson [Reuters] Journal Citation Reports (JCR) IF ≥ 0 Papers in internationally recognised journals Papers in other peer-reviewed journals Other papers, etc.	25** 15 5 2	1 st level	National expert evaluation of papers presented by institutions as highest level	1–10 score	National expert evaluation of papers presented by institutions (proportional to researchers' number)	1–5 score	National expert evaluation of papers presented by institutions (proportional to researchers' number)	1–10 score
						2 nd level	Papers in peer-reviewed journals	15 points	Papers in peer-reviewed journals & book chapters	3 points	Thomson Reuters JCR IF ≥ 0 Papers in peer-reviewed journals & book chapters	3** points 2 points
							Other papers, etc.	5 points	Other papers, etc.	1–2 points	Other papers, etc.	1 point
Papers in other peer-reviewed journals	10 (12*)	Other papers, etc.	2–4	Other papers, etc.	2							
Other papers	4 (5*)											

– in research on Lithuanistics; * calculation by formula (1) ** calculation by formula (2)

Table 2. Shift in criteria used for Lithuanian research assessment of research papers in Physical, Biomedical and Technological Sciences (according to Lithuanian science classification).

Assessment categories	2005		2006 and 2008		2009		2010; 2011		2015	
	Req. for a journal	Value, points	Requirements for a journal	Value, points	Requirements for a journal	Value, points	Requirements for a journal	Value	Requirements for a journal	Value
A-category papers 1 st level	Thomson ISI Master Journal List	10	Thomson [Reuters] Journal Citation Reports (JCR) IF ≥ 0	30**	Thomson Reuters JCR with IF > 20% AIF	15**	National expert evaluation of papers presented by institutions (proportional to researchers' number)	1–5 score	National expert evaluation of papers presented by institutions (proportional to researchers' number)	1–5 score
B-category papers (% of A-cat.) 2 nd level	–	–	Physical sciences: B ≤ 0.1 A Biomedicine: B ≤ 0.2 A Technologies: B ≤ 0.3 A	6	Physical sciences: B ≤ 0.2 A Biomedicine: B ≤ 0.2 A Technologies: B ≤ 0.3 A	5	Thomson Reuters JCR with: (1) IF > 20% AIF; (2) 20% citations from journals with IF > AIF	3** points	Thomson Reuters JCR with IF > 20% AIF	3** points

paper published in any publication cited at least 90 times by journals listed in ISI the Master Journal List. Those citations are calculated since 1990 only. ** Calculation by formula (2).

Research-driven Classification and Ranking in Higher Education: An Empirical Appraisal of a Romanian Policy Experience

Gabriel-Alexandru Vîiu¹, Mihai Păunescu², and Adrian Miroiu³

¹*gabriel.alexandru.viiu@snsa.ro*, ²*paunescu.mihai@gmail.com*, ³*admiroiu@snsa.ro*

National School of Political and Administrative Studies, Povernei Street 6, 010643 Bucharest (Romania)

Abstract

In this paper we investigate the problem of university classification and its relation to ranking practices in the policy context of an official evaluation of Romanian higher education institutions and their study programs. We first discuss the importance of research in the government-endorsed assessment process and analyze the evaluation methodology and the results it produced. Based on official documents and data we show that the Romanian classification of universities was implicitly hierarchical in its conception and therefore also produced hierarchical results due to its close association with the ranking of study programs and its heavy reliance on research outputs. Then, using a distinct data set on research performance we further explore the differences between university categories. We find that our alternative assessment of research productivity – measured with the aid of Egghe's g-index – only provides empirical support for a dichotomous classification of institutions.

Conference Topic

University Policy and Institutional Rankings

Introduction

Since the beginning of the 1980s nationally relevant university research coupled with the pressure for accountability have increasingly shaped the policies and priorities of individual universities (Geuna, 2001). Since then, the growing importance of research has been continually underscored by transnational policy documents such as the EU 2020 Strategy, by implementation of performance-based research funding mechanisms which create new competitive pressures within national university systems (Hicks, 2012) and, perhaps most visibly and controversially, by national and international university rankings which fuel debates surrounding 'world-class universities' (Sadlak & Liu, 2007; Salmi, 2009; Shin & Kehm, 2013). It is now well established that "international rankings of universities have become both popular with the public and increasingly important for academic institutions" (Buela-Casal et al., 2007, p. 351). At the same time rankings have also become "successful as an agenda-setting device for both politicians and for the higher education sector" (Stensaker & Gornitzka, 2009, p. 132).

In this paper we present an empirical exploration of the research-driven ranking and classification processes directed toward the Romanian higher education institutions (henceforth "HEIs") in the policy context of a new Law on National Education. In accordance with the new law a comprehensive process of evaluation was conducted in Romania in 2011 with the dual aim of (1) classifying HEIs (at the global, institutional level) and (2) ranking their constituent study programs. The ranking and classification were conducted using a common methodology that heavily emphasized the research productivity of university staff. Our primary objective is to contribute to a better understanding of the relation between the classification and ranking processes by discussing the methodological outline of the official evaluation and by analyzing its results. To achieve this goal we rely on official documents and on data collected with regard to the actual results of the classification and ranking processes. A secondary objective of our paper is to investigate the consistency of the institutional classification categories used in the official evaluation. To do this we employ an alternative data set on research performance, measured using the g-index which – for the set of papers of an individual researcher – represents "the largest rank (where papers are arranged in

decreasing order of the number of citations they received) such that the first g papers have (together) at least g^2 citations” (Egghe, 2006, p. 144). Our goal is to investigate whether an alternative assessment of research based on this index confirms the official classification of institutions, which was largely determined by research performance.

Background

Theoretical considerations

Higher education in recent years has witnessed the emergence of numerous university rankings, which have been the focus of comprehensive studies that aimed to investigate their methodological underpinnings, theoretical outlook and practical consequences (e.g., Dill & Soo, 2005; Salmi & Saroyan, 2007; Usher & Medow, 2009; Rauhvargers, 2011). In a more recent study Hazelkorn (2013) noted no less than 10 global rankings and at least 60 countries that have introduced national rankings. All these studies highlight (among other aspects) the fundamental importance that ranking systems generally attach to research performance, the deleterious consequences that rankings may have for institutional diversity and quality and, perhaps most importantly, the methodological caution which should be exercised when undertaking and interpreting rankings.

As more and more rankings have been developed over the years and as concerns have mounted regarding their implications and methodological problems (e.g.: van Raan, 2005; Billaut, Bouyssou & Vincke, 2010; Longden, 2011), the adjacent subject of university classification has also received increased attention (see for example Shin, 2009). This has been the case especially at the broader European level where the international ranking impetus has been critically received by scholars and policymakers and carried forward in a new direction with the introduction of the U-Map and U-Multirank initiatives, which, unlike pre-existing commercial rankings, focus on a user-driven approach and emphasize multidimensionality in evaluation.

Classification of universities has tended to be a much less debated subject than rankings, but these two distinct processes are nonetheless naturally interwoven with each other. On the one hand, due to strictures of comparability “classification is a prerequisite for sensible rankings” (van der Wende, 2008, p. 49). On the other hand, classifications are often interpreted as rankings even though this is clearly against the intentions of the classifying agency. Shulman (2005) and McCormick (2008) provide several examples of how the Carnegie Classification of US HEIs is actually understood as a form of ranking by several types of stakeholders.

A useful analytical distinction made between classifications and rankings involves conceptualizing them in the context of the broader notion of institutional diversity which itself may be divided into vertical diversity and horizontal diversity. According to van Vught (2009), the former refers to differences between higher education institutions owing to prestige and reputation while the latter stems from differences in institutional missions and profiles. In light of this distinction, classifications are “eminently suited to address horizontal diversity” (van Vught & Ziegele, 2011, p. 25) while rankings “are instruments to display vertical diversity in terms of performance by using quantitative indicators” (Kaiser, Faber & Jongbloed, 2012, p. 888).

The Romanian policy of classification and ranking

In 2011, following the provisions of the new law on national education a comprehensive national evaluation was conducted for the first time by the Romanian Ministry of Education with the aim of classifying all accredited HEIs and, additionally, of ranking all accredited study programs offered by the universities. This process was by far the most elaborate evaluation of the Romanian system of higher education and the first one to explicitly

undertake an official classification of HEIs and an official ranking of their study programs on the basis of quantitative indicators.

With regard to the classification process the law stipulated that all universities must be classified as belonging to one of the following three classes: *A – universities focused on education*; *B – universities focused on education and research*; and *C – universities focused on advanced research and education*. This would point toward a functional differentiation with regard to research capacity but the law also stipulated that the allocation of public funding was to be a function of the results of the classification process: universities from class A could only receive public funding for study programs at the bachelor level, those from class B could receive funding for programs at both bachelor and master level, while those from class C were the only ones to receive public funding for all types of programs (including PhD). With regard to the ranking of study programs, the law on education did not contain any detailed provisions. However, a subsequent government decision (789/03.08.2011) established five distinct hierarchical classes A (high quality), B, C, D and E (poor quality). These program ranking classes should not be confused with the university classes.

A detailed methodology for the classification and ranking processes was made public through Ministry of Education Order 5212/ 26.08.2012. This methodology outlined a complex system of criteria, performance indicators, variables and weights. Table 1 provides a simplified account of the evaluation methodology for the particular case of social sciences. At the most general level, four common criteria were used for both classification and ranking purposes: (1) research; (2) teaching; (3) relation to the external environment; and (4) institutional capacity. The most important aspect in the evaluation process was the research performance of the staff working in the universities and/or the study programs under assessment. This is especially significant for our later use of the *g*-index.

Table 1. Criteria, indicators and weights used in the evaluation process for university classification and study program ranking (social sciences).

<i>Criteria and global weights</i>	<i>Performance indicators and weights within criterion</i>	<i>Variables within indicator</i>
I. Research (weight: 0.50)	Results of scientific research - 0.75	11
	Research funding - 0.10	5
	International recognition - 0.02	2
	PhD programs - 0.13	2
II. Teaching (weight: 0.25)	/	6
III. Relation to external environment (weight: 0.20)	Relation to economic environment - 0.20	2
	Relation to social environment - 0.05	3
	Community development - 0.45	3
	Internationalization - 0.30	9
IV. Institutional capacity (weight: 0.05)	Indicator 1 - 0.34	3
	Indicator 2 - 0.11	3
	Indicator 3 - 0.11	4
	Indicator 4 - 0.11	4
	Indicator 5 - 0.11	4
	Indicator 6 - 0.11	1
	Indicator 7 - 0.11	5

Source: Ministry of Education Order 5212/ 26.08.2012

Within the research criterion four distinct performance indicators were defined but the most important of these four was an indicator dealing with the research output of the staff members employed by the universities. This indicator had a weight of 0.75 while the other three indicators (research funding, international recognition, and PhD programs) had much lower weights (0.10, 0.02, and 0.13). This indicator of research output was itself further broken down into 11 different variables such as the relative influence score of articles, the number of publications in journals indexed in the ISI Web of Knowledge, books, book chapters, etc.

For the ranking of study programs each university reported specific data for all of the distinct programs it operated; then, global indicators were calculated at the level of the study program for the first three criteria listed in Table 1. A separate global indicator was calculated at the university level for the institutional capacity criterion. A further step then entailed the calculation of an overall *aggregated index of ranking* (AIR) based on the four global performance indicators and their attached weights. As a final step in the ranking of a study program, its AIR was compared to the highest one obtained among all the similar study programs and, based on certain predefined intervals, it was finally designated as belonging to one of the five ranking classes.

For purposes of classification a separate *aggregated index of classification* (AIC) was calculated at the global level of each university. The AICs were calculated following a formula which incorporated three factors: (1) the absolute value of the research score obtained at the global level of the HEI; (2) a more complex factor calculated as a sum of the global indicators obtained by each of the study programs organized by the HEIs; and (3) an indicator based on the confidence level given to the HEIs by the Romanian Agency of Quality Assurance in Higher Education (ARACIS) following its periodic evaluations.

Upon calculation of the AICs of all universities the class of a particular HEI could finally be determined. Similar to the process used to establish the ranking classes of study programs, in order to determine a university's class its AIC was compared to the highest one obtained within its category (comprehensive universities were compared to other comprehensive institutions, specialized HEIs were only compared to their counterparts). First, universities were sorted in descending order of their AIC scores. Then, again following predefined intervals, universities were classified in one of the three categories A, B or C.

Without going into further details, it is apparent from even a brief analysis of the methodological outline that the evaluation conducted for purposes of classification actually had the general underpinning of a ranking. This is primarily a consequence of the fact that the classification was based on the composite scores of university performance (the AICs), which were sorted in descending order and clustered in accordance with predefined thresholds. Moreover, the classification relied on the research scores obtained by the constituent study programs of the universities and, therefore, on the partial results of the ranking process of these programs. In effect, research was the object of double counting, once at the individual level of the study programs and once more at the aggregated level of the HEIs. Based only on the analysis of the methodology used in 2011, we may argue that the entire classification process was actually hierarchical in nature and that vertical, not horizontal differentiation was a foreseeable consequence not only at the level of study programs (where ranking was explicit) but also with regard to the more general level of universities (where ranking was disavowed in favour of the more neutral label of 'classification'). However, no empirical analysis has so far been undertaken with regard to the relation between the actual results of the classification and the results of the program rankings. In addition, no independent empirical test of the three classification categories has been conducted, either relying on the performance indicators initially used by the Ministry, or on alternative measures of research performance. In the following paragraphs we will address both issues in an attempt to answer several questions related to the classification and ranking processes.

Research questions

Given the unique nature of the classification and ranking processes undertaken by the Romanian Ministry of Education several important aspects invite questioning and empirical study. We will confine our analyses to the following:

1. Did the overlap in methodology with the program rankings have empirically discernible consequences for the more general process of classification? Is there a significant degree of association between particular classes of universities and particular classes of study programs? If so, which types of programs are more common in which types of university?
2. Since the classification process relied heavily on research outputs, can an alternative assessment of the research productivity of universities confirm the threefold classification? Are there significant differences with regard to the research productivity of faculty members *between* the three university classes? Furthermore, are there significant differences with regard to the research productivity of faculty members *within* the three university classes?

The first set of questions addresses the official classification and ranking processes in tandem and implies an investigation of data on the official results. The second set of questions only addresses the classification process and will be explored using a distinct approach, which will be described in the subsequent section.

Methodology

In order to investigate our first set of research questions we created a comprehensive data set of the results of the ranking process for all the study programs evaluated in 2011. We then added the results of the classification of universities in order to obtain a final data set comprising all the study programs, the ranking class in which they were placed following the evaluation process and the class in which the university managing them was placed following the separate evaluation for classification. This primary data set contains 1056 observations of distinct study programs. To test for the level of association between ranking and classification results we created contingency tables for the occurrence of particular types of study programs (i.e. ranked in class A, B, C, D, E) in the three classes of universities (i.e. class A, class B and class C). Additionally, a chi-squared test was also used to investigate the association between the classification and ranking categories.

To explore the second set of research questions we used a distinct data set composed of information on 1,2385 Romanian faculty members active in the fields of political science, sociology and marketing. Specifically, we used their g-index to conduct an alternative assessment of university research output. These staff members represent the full populations of staff employed in political science, marketing and sociology study programs and they are spread out across 64 departments (study programs) and 34 distinct universities. Information on the identity of the staff members was obtained from ARACIS and, for each of the staff members in this second data set, the g-index was extracted using Anne Harzing's Publish or Perish software (Harzing, 2007) using a procedure previously employed in Vîiu et al. (2012) in an examination of political science departments. With regard to this secondary data set, the results of the official classification of Romanian HEIs would imply that there are significant differences between the staff employed in the three university classes with respect to their research output. To test this we employ analysis of variance and subsequent Tukey HSD tests to reveal the instances where differences between g-indices are significant. We first compare the university classes globally, and then refine our analysis to take into account more granular differences between staff types. We thus compare the four staff types – assistants, lecturers, associate professors and full professors – across the three university classes in order to determine whether or not there is a structural difference between these classes.

Results and Discussion

Relation between official ranking and official classification results

With regard to our first set of research questions a review of Table 2 and Figure 1 indicates that universities classified as being focused on education have a limited number of top-performing study programs (90 ranked in A and B, i.e. 17% of all study programs in this class of universities) but cluster the most programs with middle and low performance (those ranked in classes C, D and E add up to 83% of programs managed within the universities focused on education). On the other hand, universities focused on advanced education and research hold a total of 185 study programs and 121 of these (over 65%) are ranked in class A. Another 39 are ranked in class B (thus, over 86% of the programs in this class of universities are ranked in classes A and B) and only less than 5% belong to the lower performing classes D and E. Universities classified as being focused on both education and research have mixed results: out of a total of 344 study programs managed by these universities 189 (55%) are ranked in classes A and B, 28% are in class C and the remaining 17% are ranked in C, D and E.

Table 2. Contingency table of ranking classes of study programs and university classes.

<i>University class</i>	<i>A - Education</i>	<i>B - Education and research</i>	<i>C - Advanced research</i>	<i>Row total</i>
<i>Class of study program in official ranking</i>				
<i>A</i>	22 4.17%	60 17.44%	121 65.41%	203
<i>B</i>	68 12.90%	129 37.5%	39 21.08%	236
<i>C</i>	147 27.90%	97 28.20%	17 9.19%	261
<i>D</i>	112 21.25%	16 4.65%	3 1.62%	131
<i>E</i>	178 33.78%	42 12.21%	5 2.70%	225
<i>Column Total</i>	527 100%	344 100%	185 100%	1056
<i>Chi-Square test of ranking classes of study programs and university classes</i>				
	Value	df	Asymp. Sig. (2-sided)	
Pearson Chi-Square	495.433	8	.000	
N of Valid Cases	1056			

A more detailed study of the relationship between observed and expected count values of the different classes of study programs within each of the three university classes is also instructive. This study indicates a negative association between programs ranked in classes A and B and universities from class A. A further negative association can also be observed with regard to programs ranked in classes A, D, and E and universities from class B. Finally, universities from class C are negatively associated with study programs ranked in classes B, C, D, and E. On the other hand, a positive association exists between universities from class A and study programs ranked in classes C, D and E. A further positive association exists between universities from class B and programs ranked in classes B and C. Universities from class C are positively associated only with programs ranked in class A.

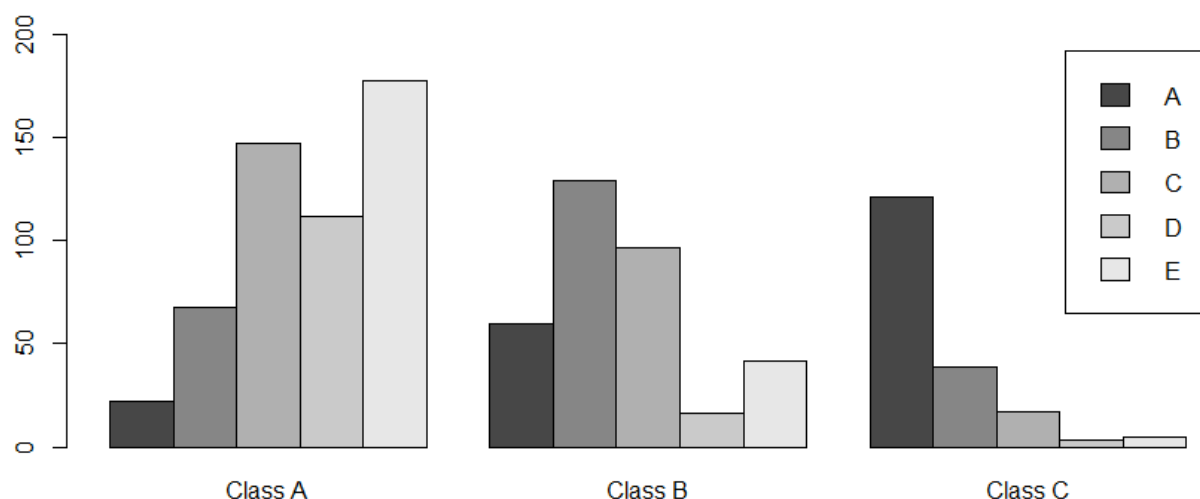


Figure 1. Distribution of study program types across the three university classes.

The results of this analysis paint a rather clear and polarized picture in which universities focused on education generally cluster study programs with poor performance while universities focused on advanced research cluster the programs with high performance. In addition, universities focused on advanced research are fewer and more selective (accounting for a total of only 185 study programs) as compared to universities focused on education (which manage a total of 527 programs). A certain hierarchy is implicit: universities focused on advanced research seem to be ‘better’ than those focused on both education and research which, in turn, are ‘better’ than those focused solely on education. However, as we mentioned earlier, these results were to be expected since both the classification and the ranking evaluation relied on a common methodology, which was mostly concerned with research performance. This leads us to our second set of research questions.

Differences in research productivity across and within university classes

We now move to explore whether our secondary data set enables us to distinguish between three university classes. In particular, what we want to see is whether the average g-index of all academic staff in class A universities is significantly lower than the average g-index of staff in class C universities and also in class B universities. The ANOVA procedure yields the results presented in Table 3. The subsequent Tukey HSD test indicates significant differences between all three means (although the confidence level for the class A – class B distinction is lower, but still above 95%) and therefore seems to provide empirical ground for the threefold classification, which was legally mandated in 2011.

Table 3. ANOVA of g-index with regard to university class (N=1,385).

<i>Model summary for ANOVA of g-index with regard to university class</i>				
	Sum of Squares	df	Mean Square	F
Between Groups	953	2	476.3	81.62
Within Groups	8065	1382	5.8	Sig.
Total	9018	1384		0.000
<i>Tukey HSD values for ANOVA of g-index with regard to university class</i>				
Comparison	Difference	Lower bound	Upper bound	p-value
Class A – Class C	-2.119	-2.513	-1.725	0.000
Class B – Class C	-1.714	-2.136	-1.293	0.000
Class B – Class A	0.405	0.054	0.756	0.019

However, the results presented in Table 3 only provide information on the global differences between university classes with regard to the *g*-indices of their entire staff, without further consideration of academic titles. Therefore, in order to test the consistency of the threefold model of classification imposed by the 2011 law, we must explore in greater depth the differences between universities, taking into account not only their classes, but also more granular differences between their academic staff. We thus set out to test not only the global aggregate differences, but also the *structural* patterns of the three classes of universities, taking into account the academic titles of the teaching staff.

In other words, bearing in mind the results of the official evaluation from 2011, we wish to know whether, for example, associate professors from class A universities are significantly different from associate professors in class B universities and from those belonging to class C and, still further, if the associate professors from class B institutions are different from those from class C. Similarly, we also wish to know whether assistants, lecturers and full professors from one class of universities are different from those belonging to the other two classes of universities. Based on such analyses we may draw more general conclusions regarding the degree of structural differentiation that exists between the three classes of universities.

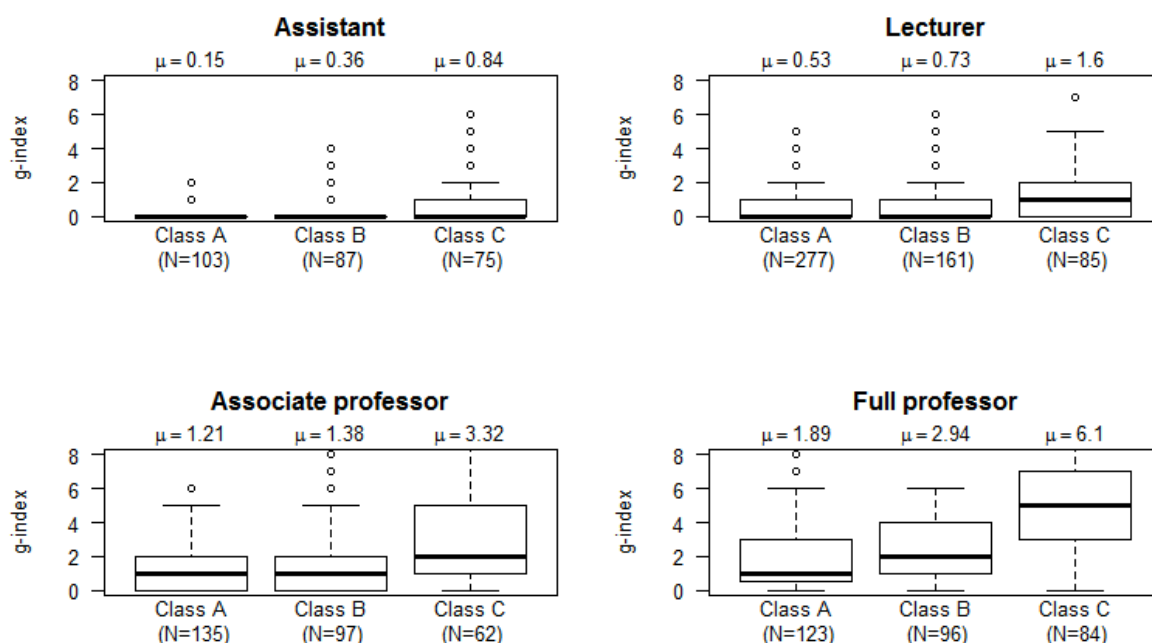


Figure 2. Distribution of *g* indices by academic title and university class.

Figure 2 illustrates the distribution of academic staff in our secondary dataset with respect to academic titles and also with regard to the class of university they belong to. Mean values are presented in the upper sections as μ . An initial visual inspection of the data would seem to indicate that in the case of assistants, lecturers and even associate professors there are no substantial differences between class A universities and those from class B. On the other hand, all three staff types working in class C universities seem to have substantially different *g*-indices compared to the ones from both class A and class B universities. A somewhat more nuanced picture emerges when looking at full professors. In this case the *g*-indices are more easily distinguishable between university classes and there indeed seem to be differences not only between class C and the other two university classes, but also between these two.

Based on the information contained in Figure 2 and on the ANOVA procedures presented in Appendix 1 we may now answer our secondary research questions. In the case of all staff members (be they assistants, lecturers, associate or even full professors) the parametric

statistical procedures show that universities classified within the official evaluation of 2011 as focused on advanced research (class C) are indeed significantly different from the other two types. In other words, assistants, lecturers, associate and full professors working in these universities focused on advanced research have significantly higher *g*-indices than their counterparts from education-centred universities, as well as from those in universities focused on both research and education. Beyond the clear distinction of staff members working in class C universities, statistical procedures also confirm something that Figure 2 reveals in a more intuitive manner: virtually no statistically significant distinction can be made between class A universities and the universities classified in 2011 as belonging to class B: assistant staff from class A universities are in no way significantly different from assistant staff working in class B universities, lecturers from one are in no way different from lecturers in the other and neither are associate professors. Even the apparent differences described by Figure 2 between full professors from class A universities and those from class B universities do not seem to be statistically meaningful either, as can be observed in Appendix 1. This suggests that a dichotomous classification would fit the data better than the threefold model imposed by the 2011 law.

So far we have argued that the data we have available clearly indicate significant *inter*-university differences (at least insofar as class C universities are made up of staff with higher indices than both class A and B universities). We now turn to *intra*-university differences. We have a reasonable expectation that within research universities there is a greater gap between the four staff types with regard to their scientific productivity. In other words, within class C universities we expect that the *g*-indices of assistants, lecturers, associate and full professors show greater dispersion than the corresponding indices of the equivalent staff that are employed in class A and class B universities. If we review the mean *g*-index values in Figure 2 we can observe that they appear to confirm our expectation. Whereas in the case of class A universities the gap between an average assistant and an average full professor is 1.74 and in the case of class B universities this gap is 2.58, in class C universities the difference is no less than 5.26. This indicates that full professors in research-centred universities have a substantially larger scientific contribution in their fields of study, not only when compared to staff employed in class A and class B universities, but also in comparison to their colleagues from the same university class. This suggests more competitive selection mechanisms of highly qualified academic staff in the research-centred universities compared to the other two university classes. These more competitive selection mechanisms may actually explain the institutional differences.

Concluding Remarks

The boundaries between classification and ranking of higher education institutions are often hard to establish and it is even harder to properly communicate the differences to intended stakeholders. When classification and ranking processes are carried out simultaneously and using common criteria the task of disambiguation becomes virtually impossible and the risk that a classification is perceived as a ranking increases exponentially. In the case of the evaluation conducted in Romania in 2011 the boundaries between classification and ranking were weak from the very inception of these evaluation processes in the law on education. The official methodology for classification and ranking further obscured the differences between the two due to its reliance on common criteria and indicators, most notably the research performance of academic staff employed by the HEIs.

By analysing the official methodology we have shown that the classification of Romanian HEIs carried out in 2011 had the underpinning of a ranking. By further analysing the results of both the classification and ranking processes we have shown that there is a clear association between the outcomes of the global process of classification and those of the more

specific process of program ranking: a polarized landscape thus emerges in which HEIs classified as focused on education cluster the overwhelming part of poor performing programs, while universities classified as focused on advanced research cluster the better part of the top performing programs.

The intermediate class of universities focused on both education and research presents mixed results. However, by conducting an alternative assessment of the research performance of the individual staff employed by Romanian universities in three fields of study we have shown that the threefold classification may not have a sufficiently robust empirical grounding, at least insofar as social sciences are concerned. By using the *g*-index as a concise measure of research performance we have illustrated the fact that the intermediate universities focused on both education and research may not be sufficiently distinct from the universities focused on education and therefore this intermediate class might have a certain degree of redundancy. When looking in our data set of 1385 staff members only at the aggregate results across university classes we do find some empirical grounding for the three classes defined in 2011. However, when analysing in greater detail the structure based on the academic titles and positions, we find less empirical grounds for the threefold classification as most of the staff employed in class A and class B universities are virtually indistinguishable from one another (i.e. assistants, lecturers and associate professors). It is only full professors that seem to make a more substantial difference between class A and class B universities, thus narrowly substantiating a threefold classification, which might otherwise well be a simpler dichotomous one. Previous extensive studies on the quality of Romanian higher education (Păunescu et al., 2012; Vlăsceanu et al., 2011; Miroiu & Andreescu, 2010) revealed the structural isomorphism of the Romanian higher education organizations. The undifferentiated set of standards that all institutions must comply with for purposes of accreditation and public funding led the institutions to adopt similar strategies for achieving these objectives. This is reflected in the poor differentiation and homogeneity of HEIs as shown by their similar scores in the external evaluation of the accreditation agency, similar missions, similar achievements on various performance indicators, etc. While there is empirical support for the vertical differentiation between advanced research universities (usually traditional, older universities) and the rest (more recent, including all private initiatives), the actual structures of the bulk of HEIs, including class A and class B universities, reveal more similarities than differences. These findings should of course be considered under the due caveat that our results are based only on data collected for social sciences.

Acknowledgments

Financial support from the National Research Council (grant number PN-II-ID-PCE-2011-3-0746) is gratefully acknowledged by Gabriel Viu and Adrian Miroiu.

References

- Billaut, J.-C., Bouyssou, D. & Vincke, P. (2010). Should you believe in the Shanghai ranking? An MCDM view. *Scientometrics*, 84, 237–263
- Buela-Casal, G., Gutiérrez-Martínez, O., Bermúdez-Sánchez, M.P. & Vellido-Muñoz, O. (2007). Comparative study of international academic rankings of universities, *Scientometrics*, 71, 349–365
- Dill, D. & Soo, M. (2005). Academic quality, league tables, and public policy: A crossnational analysis of university ranking systems. *Higher Education*, 49, 495–533
- Egghe, L. (2006). Theory and practise of the G-index. *Scientometrics*, 69, 131–152
- Geuna, A. (2001). The changing rationale for European university research funding: Are there negative unintended consequences? *Journal of Economic Issues*, 35, 607–632
- Harzing, A.W. (2007). Publish or Perish, available from <<http://www.harzing.com/pop.htm>>
- Hazelkorn, E. (2013). How rankings are reshaping higher education. In Climent, V., Michavila, F. & Ripolles, M. (Eds.), *Los rankings universitarios: Mitos y realidades*. Ed. Tecnos
- Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, 41, 251–261

- Kaiser, F., Faber, M. & Jongbloed, B. (2012). U-Map, university activity profiles in practice. In Curaj, A., Scott, P., Vlăsceanu, L., Wilson, L. (Eds.), *European Higher education at the Crossroads: Between the Bologna Process and National Reforms* (pp. 887–903). Dordrecht: Springer
- Longden, B. (2011). Ranking indicators and weights. In Shin, J.C., Toutkoushian, R.K. & Teichler, U. (Eds.), *University Rankings. Theoretical Basis, Methodology and Impacts on Global Higher Education* (pp. 73–104). New York: Springer
- McCormick, A. (2008). The complex interplay between classification and ranking of colleges and universities: Should the Berlin principles apply equally to classification? *Higher Education in Europe*, 33, 209–218
- Miroiu, A. & Andreescu, L. (2010). Goals and instruments of diversification in higher education. *Quality Assurance Review for Higher Education*, 2, 89–101
- Păunescu, M., Florian, B. & Hâncean, M.-G. (2012). Internalizing quality assurance in higher education: Challenges of transition in enhancing the institutional responsibility for quality. In Curaj, A., Scott, P., Vlăsceanu, L., Wilson, L. (Eds.), *European Higher education at the Crossroads: Between the Bologna Process and National Reforms* (pp. 317–338). Dordrecht: Springer.
- Rauhvargers, A. (2011). *Global University Rankings and Their Impact*. Brussels: European University Association
- Sadlak, J. & Liu, N.C. (2007). *The World-class University and Ranking: Aiming beyond Status*. Bucharest: UNESCO-CEPES
- Salmi, J. (2009). *The Challenge of Establishing World-class Universities*. Washington DC: World Bank
- Salmi, J. & Saroyan, A. (2007). League tables as policy instruments: uses and misuses. *Higher Education Management and Policy*, 19, 31–68
- Shin, J.C. (2009). Classifying higher education institutions in Korea: A performance-based approach. *Higher Education*, 57, 247–266
- Shin, J.C. & Kehm, B. (Eds.). (2013). *Institutionalization of World-class University in Global Competition*. Dordrecht: Springer
- Shulman, L.S. (2005). Classification's complexities. *The Chronicle of Higher Education* (November 11, 2005), 52, p. B20
- Stensaker, B. & Gornitzka, A. (2009). The ingredients of trust in European higher education. In Kehm, B.M., Huisman, J. and Stensaker, B. (Eds.), *The European Higher Education Area: Perspectives on a Moving Target* (pp. 125–139). Rotterdam: Sense Publishers
- Usher, A. & Medow, J. (2009). A global survey of university rankings and league tables. In Kehm, B.M. and Stensaker, B. (Eds.), *University Rankings, Diversity, and the New Landscape of Higher Education* (pp. 3–18). Rotterdam: Sense Publishers
- van der Wende, M. (2008). Rankings and classifications in higher education: A European perspective. In Smart, J. C. (Ed.) *Higher Education: Handbook of Theory and Research* (pp. 49–72), Vol. XXIII, Springer.
- van Raan, A. F. J. (2005). Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62, 133–143
- van Vught, F. (2009). Diversity and differentiation in higher education. In van Vught, F. (Ed.) *Mapping the Higher Education Landscape. Towards a European Classification of Higher Education* (pp. 1–16). Dordrecht: Springer
- van Vught, F. & Ziegele, F. (Eds.).(2011). *Design and Testing the Feasibility of a Multidimensional Global University Ranking. Final Report*. Consortium for Higher Education and Research Performance Assessment, CHERPA-Network
- Vîiu, G.-A., Vlăsceanu, M., & Miroiu, A. (2012). Ranking political science departments: the case of Romania. *Quality Assurance Review for Higher Education*, 4, 79-97
- Vlăsceanu, L., Miroiu, A., Păunescu, M. & Hâncean, M.-G. (Eds.). (2011). *Barometrul calității 2010. Starea calității în învățământul superior din România*. Brașov: Editura Universității Transilvania din Brașov.

Appendix 1. Tests of difference for g-index across academic titles and university classes.

<i>1. Model summary for ANOVA of g-index of assistant staff with regard to university class</i>				
	Sum of Squares	df	Mean Square	F
Between Groups	20.68	2	10.341	13.29
Within Groups	203.82	262	0.778	Sig.
Total	224.50	264		0.000
<i>Tukey HSD values for ANOVA of g-index of assistant staff with regard to university class</i>				
Comparison	Difference	Lower bound	Upper bound	p-value
Class A – Class B	0.212	-0.090	0.515	0.224
Class C – Class A	0.684	0.369	1.000	0.000
Class C – Class B	0.472	0.144	0.799	0.002
<i>2. Model summary for ANOVA of g-index of lecturers with regard to university class</i>				
	Sum of Squares	df	Mean Square	F
Between Groups	73.7	2	36.85	25.39
Within Groups	754.8	520	1.45	Sig.
Total	828.5			0.000
<i>Tukey HSD values for ANOVA of g-index of lecturers with regard to university class</i>				
Comparison	Difference	Lower bound	Upper bound	p-value
Class A – Class B	0.195	-0.085	0.475	0.232
Class C – Class A	1.062	0.710	1.413	0.000
Class C – Class B	0.867	0.487	1.246	0.000
<i>3. Model summary for ANOVA of g-index of associate professors with regard to university class</i>				
	Sum of Squares	df	Mean Square	F
Between Groups	204.8	2	102.40	24.44
Within Groups	1219.2	291	4.19	Sig.
Total	1424			0.000
<i>Tukey HSD values for ANOVA of g-index of associate professors with regard to university class</i>				
Comparison	Difference	Lower bound	Upper bound	p-value
Class A – Class B	0.166	-0.475	0.808	0.813
Class C – Class A	2.107	1.367	2.847	0.000
Class C – Class B	1.941	1.157	2.725	0.000
<i>4. Model summary for ANOVA of g-index of full professors with regard to university class</i>				
	Sum of Squares	df	Mean Square	F
Between Groups	914	2	457.0	34.83
Within Groups	3936	300	13.1	Sig.
Total	4850			0.000
<i>Tukey HSD values for ANOVA of g-index of full professors with regard to university class</i>				
Comparison	Difference	Lower bound	Upper bound	p-value
Class A – Class B	1.053	-0.108	2.215	0.084
Class C – Class A	4.212	3.005	5.420	0.000
Class C – Class B	3.159	1.884	4.433	0.000

Looking beyond the Italian VQR 2004-2010: Improving the Bibliometric Evaluation of Research

Alberto Anfossi^{1,2}, Alberto Ciolfi¹, Filippo Costa^{1,3}

¹albertofrancesco.anfossi@anvur.it, ¹alberto.ciolfi@anvur.it, ^{1,3}filippo.costa@anvur.it

¹ANVUR, Via Ippolito Nievo 35, 00153 Roma (Italy)

²Compagnia di San Paolo Sistema Torino, Piazza Bernini 5, Torino (Italy)

³Dipartimento Ingegneria dell'Informazione, Università di Pisa, Pisa (Italy)

Abstract

In the recent Italian Evaluation of Research Quality exercise for the period 2004-2010 (VQR), promoted by the Italian Ministry of Education and carried by the National Agency for Research Evaluation (ANVUR), metrics were massively employed. The use of Impact Factor or article citations (or both) are usually considered a powerful tool for supporting the peer review process but the replacement of the latter with an automatic evaluation tool has been always considered risky. Here we propose a possible prescription to overcome some limitations of the bibliometric evaluation carried out within the context of the VQR, while, at the same time, keeping the main distinctive features of the evaluation approach unchanged, namely, a simple evaluation tool based on the combined use of the CIT and IF variables. While maintaining the basic elements of the previous algorithm unchanged and keeping the method simple and feasible on a large scale, we argue that the main flaws and limitations can be overcome.

Conference Topic

University Policy and Institutional Rankings

Introduction

The most popular European national research evaluation is the Research Assessment Exercise (RAE) in Great Britain, which started in 1986 and has been replaced, in 2013, by a new exercise - Research Excellence Framework (REF) - where citation-based metrics were employed to inform and supplement Peer Review (PR) evaluation.

In Italy, the first evaluation exercise was carried out in 2005 by the CIVR with reference to the period 2001-2003 (VTR). The VTR was fully based on the PR evaluation method, each submitted research product being assessed by a pool of experts (Minelli et al., 2008). Some studies (Reale et al. 2007; Abramo et al., 2009; Franceschet et al., 2011) analysed the outputs of the VTR comparing peer quality opinions on papers with metrics based on the Impact Factor of the journals publishing the papers, concluding that the two evaluation methods significantly overlap. However, comparison of PR and bibliometric evaluation methodologies has been largely debated in the literature (Barker, 2007; Moed, 2006; Harnad, 2009; Norris et al. 2003, Butler et al., 2003; Bence et al., 2009, Asknes, et al. 2004) with not always concordant outcomes. The use of Impact Factor or article citations or both are usually considered a powerful tool for supporting the PR process but the replacement of the latter with an automatic evaluation tool has been always considered risky.

In the recent Italian Evaluation of Research Quality exercise for the period 2004-2010 (VQR), promoted by the Italian Ministry of Education and carried by the National Agency for Research Evaluation (ANVUR), metrics were massively employed. Around 200.000 research outputs, mainly journal articles or reviews (both called 'paper' in the following), were evaluated, of which 46,5% by use of a bibliometric algorithm (Ancaiani et al., 2014).

Bibliometric Evaluation in the VQR 2004-2010

According to the Ministerial Decree number 17 of July 15th, 2011 promoting the VQR, each paper submitted for evaluation is classified in one of four possible classes of merit, defined as follows: “Excellent” (E): when the paper falls in the top 20% of the world production in a given Subject Category (SC) and in a given year; “Good” (G): when the paper falls in the following 20%; “Acceptable” (A): when a paper falls in the following 10%; “Limited” (L): when a paper falls in the bottom 50%.

In bibliometric areas, the strategy to assign a paper to a given class was based on the combined use of two variables: (i) CIT: number of citations collected by the paper up to December 31st, 2011 and (ii) IF: Impact Factor (or equivalent indexes) of the Journal in the year of publication of the paper. Each paper was submitted by the Organization (i.e. universities or public research bodies) and then uniquely assigned to a thematic evaluation panel (called Group of Experts for Evaluation, GEV) and to a Subject Category (SC), or All Journal Science Category (ASJC) as defined by ISI Web of Knowledge® or Scopus databases, respectively. In each SC/ASJC and for each year it is possible to construct the cumulative distribution function of the two variables¹, thus assigning to each paper its CIT and IF percentile. In the VQR three thresholds for both IF and CIT were defined to distinguish among the four classes specified in the Ministerial Decree. In the space spanned by IF (x-axis) and CIT (y-axis) it is therefore possible to focus on the region $Q = [0,1] \times [0,1]$ and plot the publications distribution defined on the basis of their CIT and IF percentile (Fig. 1, where each dot represents a paper denoted by its CIT and IF percentile). Each GEV had the freedom to assign the “off-diagonal” sub-squares (blocks) of the whole region Q , identified by the intersection of the “threshold segments”, to a class of merit, thus completing the automatic phase of the evaluation process. Indeed, the diagonal blocks were quite naturally assigned to the four classes: the intersection of “top 20% for CIT” with “top 20% for IF” was straightforwardly associated to the “Excellent” class of merit, and so on. The choice to assign an off-diagonal block to a class was performed according to basically two drivers: first and foremost, the qualitative insight of the GEV on the scientific field and its publication practices (e. g. lag in citations, etc.) and second, the attempt to keep the final assignment as close as possible to the world distribution D specified in the Ministerial Decree.

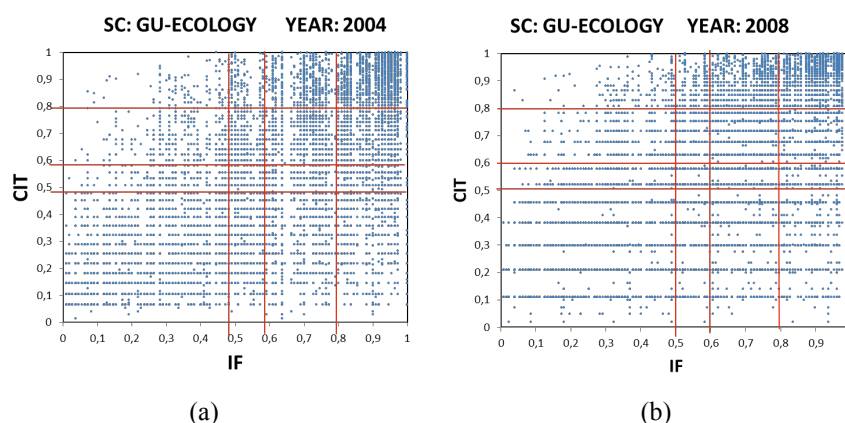


Figure 1. Papers distribution in a given SC and in two different years.

Such an approach showed some limitations that we summarize schematically:

¹ CIT: by ordering the total number of paper published in that SC and in that year in decreasing order from the highest to the lowest cited; IF: by ordering the Journals belonging to that SC in that year in decreasing order from the highest impact factor to the lowest.

Absence of “micro calibration”: all the GEVs except for GEV 02 (Physical sciences) chose a single assignment (typically, one for years 2004-2008 and one for years 2009-2010), i.e., association of blocks to classes of merit, and did not went through a micro calibration at the level of the single SC and single year. Considering that: (i) for each GEV the number of relevant SC² was typically of the order of 50 and (ii) the distribution of the papers in Q was totally not uniform and invariant, rather, it varied significantly from one SC to another and from one year to another (see for instance Fig. 4). The absence of a micro calibration affected the possibility to comply with the distribution D punctually (and not only on average).

Structure of the blocks: (i) as showed in Figure 1 the threshold segments are parallel to the x/y axis. This is not convenient given the discrete nature of the two variables under consideration. (i) It can be easily noted in the plot that the points (corresponding to papers) are distributed in rows, according for instance to the limited number of journals present in a SC. As a consequence, the evaluation may not be robust enough, in the sense that a slight perturbation in the thresholds can modify the final class allocation for whole set of papers. (ii) It is quite hard, if not impossible, to comply with the distribution D by leveraging on the sole degrees of freedom given by the possibility to assign the off-diagonal blocks to a final class of merit. In other words, the constraint of assigning to a single class an entire block is too binding and tends to move too many paper from one class of merit to an another. (iii) The degrees of freedom are even reduced by the need to avoid that two non-adjacent classes of merit (say, “Good” and “Limited”) can be adjacent in Q, as shown in Fig. 2.

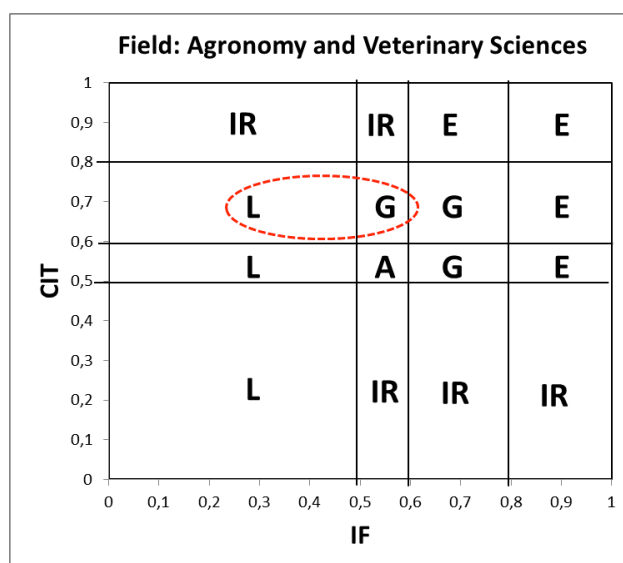


Figure 2. Algorithm used to evaluate research products in the Agronomy and Veterinary Science field: two non-adjacent classes of merit are adjacent in Q (red circle). “IR” indicates products that are left undecided by the algorithm and are eventually evaluated by peer review.

The new proposed approach

In the following we discuss a possible prescription to overcome these limitations while, at the same time, keeping the main distinctive features of the evaluation approach unchanged, namely, a simple evaluation tool based on the combined use of the CIT and IF variables. This can be done through the use of three diagonal segments with generic slope (Fig. 3).

² By relevant we mean that a great number (more than one hundred) of papers to be evaluated fell under that SC.

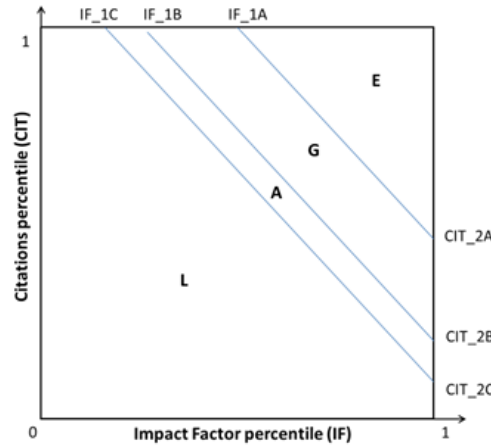


Figure 3. New prescription for combining the CIT and IF variables.

Such a new prescription builds upon three main pillars:

1. The segments identifying the thresholds are now drawn as a linear combination of the CIT/IF thresholds, thus being diagonal and no more parallel to the axes;
2. CIT/IF thresholds do not have to separately satisfy the 20-20-10-50 distribution;
3. The calibration, i.e. where to position the diagonal segments in Q in order to comply with the distribution D, is now performed at the micro level of each SC, for each year and for each GEV (according to general guidelines provided by the GEV itself and based on GEV's proficiency in the specific scientific field);

This would in turn guarantee the effectiveness and the simplicity of the whole process. In Figure 4 we apply this method to some SCs.

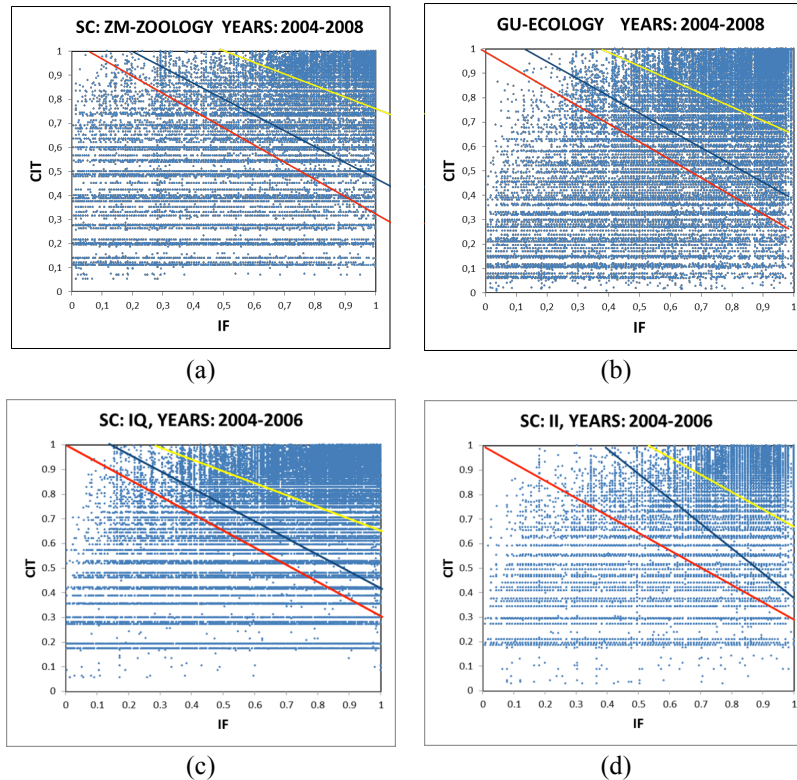


Figure 4. The application of the new algorithm in various SC and years. IQ stands for Electrical and Electronic Engineering, II stands for Engineering Chemical. The straight lines indicate the thresholds for the four classes of merit.

Comments and future developments

This new approach is characterized by a rather marked level of freedom in the choice of the position of the diagonal segments (or, equivalently, of the CIT/IF thresholds). Indeed, there is typically more than one choice that satisfies the distribution D. On the other hand one could impose additional constraints, such as for instance the parallelism between segments, based on additional empiric work and on scientific validation of the procedure (eg. by a PR comparison of the evaluation outcomes). Furthermore, such a freedom might be exploited to accommodate GEV's requirements. For instance, it would be possible to give more relevance to one of the two dimensions (IF, CIT) depending on, say, the year of publication or the citation praxes of specific disciplines (Mathematics vs Medicine being a paradigmatic example).

A significant possibility to further improve the accuracy of the method we discussed comes from a different definition of the cumulative distribution function for the IF variable. Instead of considering the number of journals belonging to a SC, one could consider the number of items (papers) published in the SC (in a given year). Actually, it is common that some journals host few thousands of items per year while other few tens or units. This induces a possible distortion that is quite evident in the plots shown below. As an example, In Figure 5 we analyze the distribution of the SC Electrical and Electronic Engineering in 2004. The distribution of the papers according to the IF and CIT percentile are depicted both considering only the number of journals in the calculation of the IF percentile and by considering also the number of item for each journal. The distributions are subdivided with different lines in order to obtain the target percentages D. It is evident that the equation of the lines is substantially different to guarantee the same final result. It is worth underlining that the lines used to subdivide the distribution reported in Figure 5(a) would result in very different percentages if applied to the distribution in Figure 5(b).

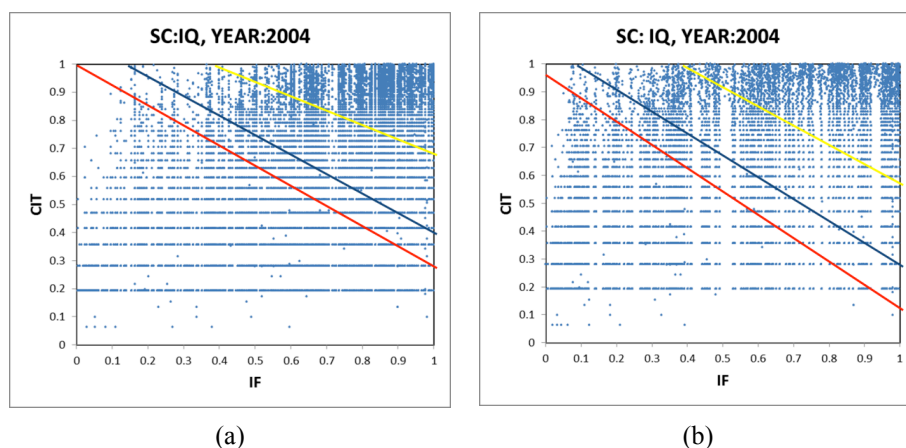


Figure 5. Distribution of the papers according to the number of journals and papers. (a) IF percentile calculated based on the number of journals (b) the IF percentile is calculated considering the number of items. The distributions are subdivided with lines in order to obtain the target percentage D.

Finally, it would be possible to improve also the CIT dimension by overcoming the concept of SC as “reference set” and move on to clustering strategies based on semantic or on citation networks. This would be more rigorous and meaningful considering the existence of a great number of journals that publish very different subjects, but it would come with a significant enhancement of the complexity of the evaluation procedure, probably not feasible for the numbers implied by a national formal evaluation, at the moment.

Results obtained so far are already highly informative about the existing strength and weakness of the Italian University research system, and provide reliable input for policy interventions. Our proposal is intended to further improve the mix of peer review and bibliometric methods through a more precise calibration of the biblio(metrics) used.

The output turns out to be rather general, thus being applicable to other national assessments based on bibliometric analysis.

References

- Ancaiani et al. (2014). Evaluating Scientific Research in Italy: the 2004-2010 Research Evaluation Exercise. *Research Evaluation*, submitted
- Minelli, E., Rebori, G., Turri, M. (2008). The structure and significance of the Italian research assessment exercise (VTR). *European Universities in Transition*. Edward Elgar Publishing.
- Reale, E., Barbara, A., & Costantini, A. (2007). Peer review for the evaluation of academic research: lessons from the Italian experience. *Research Evaluation*, 16(3), 216-228.
- Franceschet, M., & Costantini, A. (2011). The first Italian research assessment exercise: A bibliometric perspective. *Journal of Informetrics*, 5(2), 275-291.
- Abramo, G., D'Angelo, C. A., & Pugini, F. (2008). The measurement of Italian universities' research productivity by a non parametric-bibliometric methodology. *Scientometrics*, 76(2), 225-244.
- Barker, K. (2007). The UK Research Assessment Exercise: the evolution of a national research evaluation system. *Research Evaluation*, 16(1), 3-12.
- Moed, H. F. (2006). Citation analysis in research evaluation (Vol. 9). Springer.
- Harnad, S. (2009). Open access scientometrics and the UK Research Assessment Exercise, *Scientometrics*, 79(1), 147-156
- Norris, M., & Oppenheim, C. (2003). Citation counts and the Research Assessment Exercise V: Archaeology and the 2001 RAE. *Journal of Documentation*, 59(6), 709-730.
- Butler, L., & McAllister, I. (2009). Metrics or peer review? Evaluating the 2001 UK Research assessment exercise in political science. *Political Studies Review*, 7(1), 3-17.
- Bence, V., & Oppenheim, C. (2004). The influence of peer review on the research assessment exercise. *Journal of Information Science*, 30(4), 347-368.
- Asknes, D. W., Taxt, R. E. (2004). Peer reviews and bibliometric indicators: a comparative study at a Norwegian university. *Research evaluation*, 13(1), 33-41.

High Fluctuations of THES-Ranking Results in Lower Scoring Universities

Johannes Sorz¹, Martin Fieder², Bernard Wallner² and Horst Seidler²

johannes.sorz@univie.ac.at

¹University of Vienna, Office of the Rectorate, Universitätsring 1, A-1010 Vienna (Austria)

martin.fieder@univie.ac.at, wallner@univie.ac.at, horst.seidler@univie.ac.at

²University of Vienna, Department for Anthropology, Althanstrasse 14, A-1090 Vienna (Austria)

Abstract

A regression analysis of results from the Times Higher Education World University Rankings (THES-Ranking) from 2010-2014 shows high fluctuations in the rank and score for lower scoring universities (below position 50) which lead to inconsistent “up and downs” in the total results. We conclude that these fluctuations do not correspond to actual university performance. They create the impression of the THES-Ranking as a “gamble” for universities below rank 50. We suggest that THE alters its ranking procedure insofar as universities below position 50 should be ranked summarized only in groups of 25 or 50. Additionally, we argue for introducing a standardization process for THES-Ranking data by using common suitable reference data to create calibration curves represented by non-linearity or linearity.

Conference Topic

University Policy and Institutional Rankings

Introduction

Global higher education rankings have received much attention recently and, as can be witnessed by the growing number of rankings being published every year, this attention is not likely to subside. Besides the arguable use of results from global rankings as an instrument for rational university management, they remain influential for stakeholders inside and outside academia. A plethora of regional and national rankings exist, and 10 global higher education rankings are currently attempting to rank academic institutions worldwide. Numerous studies have analyzed and criticized higher education rankings and their methodologies (van Raan, 2005; Buela-Casal et al., 2007; Ioannides et al., 2007; Hazelkorn, 2007; Aguillo et al., 2010; Benito and Romera, 2011; Hazelkorn, 2011; Rauhvargers, 2011; Tofallis, 2011; Saisana et al. 2011; Safon, 2013; Rauhvargers, 2013). This casts justified doubt on a sensible comparison of universities hailing from different higher education systems and varying in size, mission and endowment based on mono-dimensional rankings and league tables (Hazelkorn, 2014). Several studies have demonstrated that data used to calculate ranking scores can be inconsistent. Thus, bibliometric data from international databases (Web of Science, Scopus), used in most global rankings to calculate research output indicators, favor universities from English-speaking countries and institutions with a narrow focus on highly-cited fields, which are well covered in

these databases. This puts universities from non-English-speaking countries, with a focus on the arts, humanities and social sciences, at a disadvantage when being compared in global rankings (Calero-Medina et al., 2008; van Raan et al., 2011; Waltman et al., 2012). Data submitted by universities to ranking agencies (e.g. personnel data, student numbers) can be problematic to compare due to different standards. These incompatibilities are being amplified because university managers have become increasingly aware of global rankings and try to boost their performance by “tweaking” the data they submit to the ranking agencies (Spiegel Online, 2014). Beyond all the data issues, there is the effect that universities with lower positions in the rankings often encounter volatile ups and downs in their consecutive year-to-year ranks. This creates the sensation of contending in a “gamble” in which results are calculated at random by ranking agencies. Such effects make global university rankings in many cases an inappropriate tool for university managers: the ranking results simply do not reflect the universities’ actual performance or their management strategies. Volatile jumps are also difficult to explain to the media, which often engage in sensationalism when covering rankings by interpreting subtle changes of scores, even within the margins of statistical deviations, as substantial shifts in performance. Bookstein et al. (2010) found unacceptably high year-to-year variances in the score of lower ranked universities caused by statistical noise in the Times Higher Education World University Ranking (THES), one of the currently most popular global rankings. We again observed puzzling variances in the THES-Ranking 2014-2015, published in October 2014. Accordingly, we here analyze the fluctuations in score and rank of the THES-Ranking by calculating a regression analysis for consecutive years for 2010-2014 to determine the random component of these fluctuations. The methodology of the THES-Ranking was revised several times in varying scale, before and after the split with Quacquarelli Symonds (QS) in 2010 and the new partnership with Thompson Reuters. Times Higher Education (THE) calculates 13 performance indicators, grouped into the five areas Teaching (30%), Research (30%), Citations (30%), Industry income (2.5%) and International outlook (7.5%). However, THE does not publish the scores of individual indicators, only those of all five areas combined. Since 2010, the research output indicators are calculated based on Web of Science data. Most of the weight in the overall score is made up by the normalized average citations per published paper (30%), and by the results of an academic reputation survey (33%) assessing teaching and research reputation and influencing the scores of both areas (Rauhvargers, 2013; THE, 2014). In the past, criticism has been levied against this survey. Academic peers can choose universities in their field from a preselected list of institutions and, although universities can be added to the list, those present on the original list are more likely to be nominated. This leads to a distribution skewed in favor of the institutions at the top of the rankings (Rauhvargers, 2011; Rauhvargers, 2013). THE allegedly addressed this issue by adding an exponential component to increase differentiation between institutions, yet no information is available on its mode of calculation (Baty, 2011; Baty, 2012).

Methods

We used the publicly available data on scores and ranks from the THES-Ranking for the years 2010, 2011, 2012, 2013 and 2014, including only those universities ranked from 1 to 200. We performed the following analysis: i) we regressed the scores of the ranking of the year $t-1$ on the scores of the year t ; ii) we regressed the ranks of the ranking of the year $t-1$ on the ranks of the year t ; iii) we plotted the scores in descending order and iv) we determined the random component of the fluctuations in the ranks from year to year.

Results

Regression of the scores and ranks of two consecutive years

The regression of the scores—particularly of the ranking 2010-2011 regressing on the scores of the ranking of 2011-2012—shows a very high fluctuation/noise (Figure 1a), especially for the lower ranked universities. Moreover, the noise among the lower ranked universities seems to be higher compared to the already very noisy THES-Ranking performed by QS before 2010 (Bookstein et al., 2010, Figure 1). Note that in the rankings in the years following 2010-2011, the noise in the THES-Ranking did improve (Figure 1b-d).

Association between Scores and Ranks

Nonetheless, a general problem of the THES-Ranking remains: the difference in the scores among the 50 highest scoring universities is considerably higher compared to the difference among the lower scoring universities. This clearly suggests a non-linear relationship between scores and ranks (Figure 2 a-e). The consequence is that the ranks of the high scoring universities are much more robust to deviations in the scores from year to year. In the lower ranking universities, however, even very small, more or less random deviations (around 0.5%) lead to unexpected “high jumps” in the ranks from year to year (Figure 1e-h).

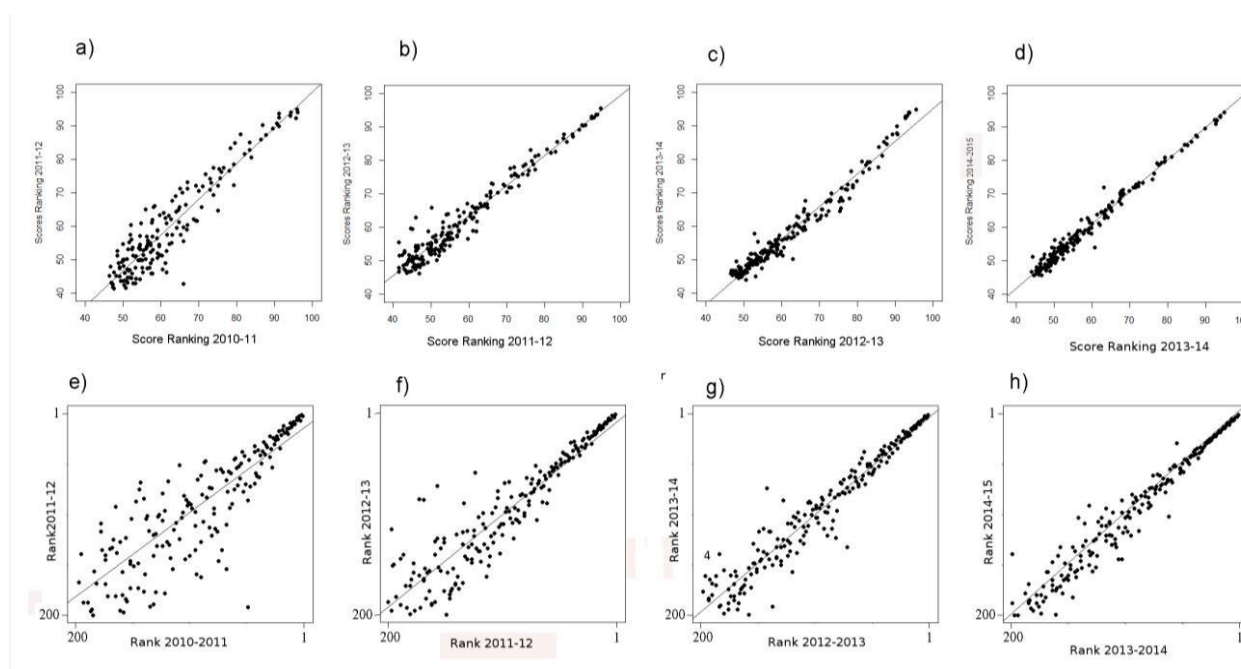


Figure 1a-1d) Scores of the year t-1 regressing on the score of the year t from the ranking 2010-11 on. Figure 1e-1h) Ranks of the year t-1 regressing on the ranks of the year t from the ranking 2010-11 on. Linear regression line indicates perfect association, e.g. no changes in ranks and scores between two consecutive rankings.

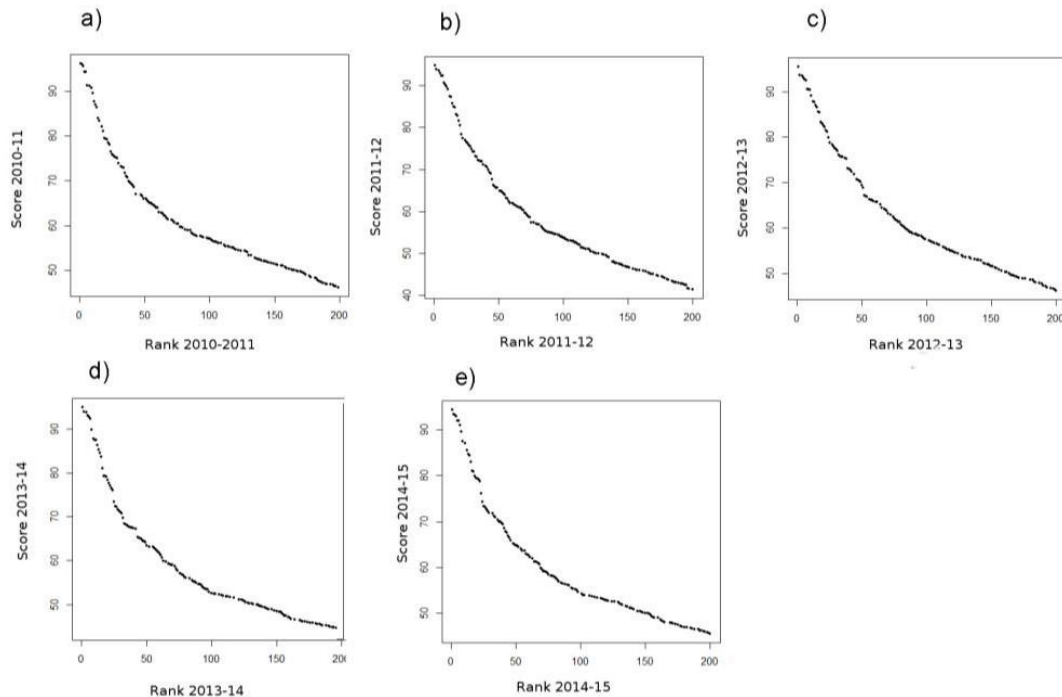


Figure 2 a-e). Ranks plotted against scores for the THES-Ranking a) 2010-11; b) 2011-12; c) 2012-13; d) 2013-2014; e) 2014-15

Discussion and Outlook

High ranking positions achieved by a small group of universities are often self-perpetuating, especially due to the intensive use of peer review indicators, which improve chances of maintaining a high position for universities already near the top (Bowman & Bastedo, 2011; Rauhvargers, 2011). This phenomenon also corresponds to the Matthew effect, which was coined by Merton (1968) to describe how eminent scientists will often get more credit than a comparatively unknown researcher, even if their work is similar: credit will usually be given to researchers who are already famous. The intensive and exaggerated discussion in the media of the “up and downs” of universities in the THES-Ranking is particularly misleading for the lower scoring universities (below approximately a score of 65% and a rank of 50; above scores of 65%, the relationship between ranks and scores is steeper, and it flattens for scores below 65%). This is because the ranking positions suggest substantial shifts in university performance despite only very subtle changes in score. In fact, merely random deviations must be assumed. One reason lies in the weighing of indicators by THE, with the emphasis on citations and peer review (totaling more than 65% of the total score). For lower ranked universities, a few highly cited publications, or the lack thereof, or few points asserted by peers in the reputation survey, probably make a significant difference in total score and position. In a follow up study that is currently under review we compared the results from THES with the results of the ARWU-Ranking (aka Shanghai-Ranking). Although the ARWU-Ranking seems to be more robust than the THES-Ranking (less year-to-year fluctuations probably due to the omittance of peer review indicators), we also found fluctuations below rank 50 and patterns of non-linearity between ranks and scores. Furthermore we found out that year-to-year results do not correspond in THES- and ARWU-Rankings for universities below that rank.

Ranking results have a major influence on the public image of universities and even impact their claim to resources (Espeland & Saunder, 2007; Hazelkorn, 2011). Accordingly, such fluctuations in the THES-Rankings can have serious implications for universities, especially when the media or stakeholders interpret them as direct results of more or less successful

university management. Our initial data in combination with the data from the literature strongly suggests that universities as well as policy makers and stakeholders should avoid to use rankings, especially league-tables, for management purposes or for strategic planning.

More specifically, the THES-Rankings in their current form have very limited value for the management of universities ranked below 50. This is because the described fluctuations in rank and score probably do not reflect actual performance, whereby the results cannot be used to assess the impact of long-term strategies. Thus, results from the THES (and to some extent also the ARWU) should be used only with great discretion. The low correlation between the ranks of the THES and the ARWU ranking, particularly for the universities ranked below 50 in both rankings, creates another serious doubt if rankings should be used for any management purposes at all. Maybe a “meta-analysis” of rankings could be reasonable to derivate consistent and reliable results from rankings. If done, such a meta-analysis should include as many rankings as possible to reduce random perturbations.

Multidimensional rankings, like the U-Multirank (<http://www.u-multirank.eu>), seem to offer a more versatile picture that reflects both the diversity of higher education institutions and the variety of dimensions of university excellence, allowing university managers to compare institutions on various levels. Although multidimensional rankings do get less public attention than league-tables and they can be prone for errors for the same reasons as monodimensional rankings (e.g., incompatibility of data provided by the universities), from the perspective of a university manager, they offer a more diverse toolset to gauge an institutions strength and weaknesses and to benchmark comparable universities.

“Rankings are here to stay, and it is therefore worth the time and effort to get them right,” warns Gilbert (2007). That is especially true for monodimensional rankings, like the THES, that spark a lot of media attention. What could be done to address the fluctuations in the THES-Rankings for universities below rank 50 and to avoid the impression of a “gamble” in which THE “rolls a dice” to determine scores and ranks? THE has already addressed fluctuations to some extent by ranking universities only down to position 200, followed by groups of 25 from 201-300 and groups of 50 from 300 to 400. Nonetheless, based on our data we believe that this is not going far enough and suggest that universities should be summarized in groups of 25 or 50 below the position of 50.

The analyzed curves of scores vs. ranking positions in Figure 2 do have analogous characteristics for example to semi-logarithmic curves produced in analytic biochemistry. The accuracy of such curves is limited to the steepest slope of the curve, whereas asymptote areas deliver higher fuzziness (Chan, 1992). Thus, a further suggestion to avoid the blurring dilemma is the methodological approach of introducing a standardization process for THES-Ranking data. This would involve using common suitable reference data to create calibration curves represented by non-linearity or linearity. However, more research in this area is necessary.

The results presented in this paper are only the starting point and we plan to do more in-depth analyses of the variations in the various indicators in the future. We already have extended our analysis to include the ARWU-Ranking (paper currently in review) and we plan to analyze and compare other major higher education rankings (e.g. the QS-Ranking) in future publications to assess their usability for university management purposes.

References

- Aguillo, I.F., Bar-Ilan, J., Levene, M. & Ortega, L.J. (2010). Comparing university rankings. *Scientometrics* 85, 243–256.
- Benito, M. & Romera, N. (2011). Improving quality assessment of composite indicators in university rankings: a case study of French and German universities of excellence. *Scientometrics* 89, 153–176.
- Baty, P. (2011, October 6) THE Global Rankings: Change for the better, *Times Higher Education*, <http://www.timeshighereducation.co.uk/world-universityrankings/2011-12/world-ranking/methodology>

- Baty, P. (2012, October 4). The essential elements in our world-leading formula, *Times Higher Education*, <http://www.timeshighereducation.co.uk/worlduniversity-rankings/2012-13/world-ranking/methodology>
- Bowman, A.M. & Bastedo, N.M. (2011). Anchoring effects in world university rankings: exploring biases in reputation scores. *Higher Education* 61, 431–444
- Bookstein, F.L., Seidler, H., Fieder, M., & Winckler, G. (2010). Too much noise in the Times Higher Education rankings. *Scientometrics* 85, 295–299.
- Buela-Casal, G., Gutierrez-Martinez, O., Bermudes-Sanchez, M., & Vadillo-Munoz, O. (2007) Comparative study of international academic rankings of universities. *Scientometrics* 71, 349–365.
- Calero-Medina, C., López-Illescas, C., Visser, M.J and Moed, H.F. (2008). Important factors when interpreting bibliometric rankings of world universities: an example from oncology. *Research Evaluation*, 17 (1), 71–81.
- Chan, D.W. (ed) (1992) *Immunoassay Automation: a Practical Guide*. San Diego, CA: Academic Press.
- Espeland, W.N & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds, *American Journal of Sociology*, 113 (1), 1–40.
- Gilbert, A. (2007). Academics strike back at spurious rankings. *Nature*, 447, 514–515.
- Hazelkorn, E., (2007) Impact and influence of league tables and ranking systems on higher education decision-making. *Higher Education Management and Policy*, 19 (2), 87–110.
- Hazelkorn, E. (2011) *Rankings and the Reshaping of Higher Education: the Battle for World Class Excellence*. Basingstoke: Palgrave-MacMillan.
- Hazelkorn, E. (2014) Reflections on a decade of global rankings: what we've learned and outstanding issues, *European Journal of Education*, 49 (1), 12–28.
- Ioannidis, J. P. A., Patsopoulos, N. A., Kavvoura, F. K., Tatsioni, A., Evangelou, E., Kouri, I., et al. (2007). International ranking system for universities and institutions: A critical appraisal. *BMC Medicine* 5, 30.
- Merton, R.K. (1968). The Matthew effect in science. *Science*, 159, 56–63.
- Rauhvargers, A. (2011) EUA Report on Global Rankings and their Impact – Report I (European University Association). http://www.eua.be/pubs/Global_University_Rankings_and_Their_Impact.pdf
- Rauhvargers, A. (2013). EUA Report on global rankings and their Impact – Report II (European University Association). http://www.eua.be/Libraries/Publications_homepage_list/EUA_Global_University_Rankings_and_Their_Impact_-_Report_II.sflb.ashx.
- Safon, V. (2013). What do global university rankings really measure? The search for the X factor and the X entity. *Scientometrics*, 97, 223–244.
- Saisana, M., d'Hombres, B., & Saltelli X. (2011). A Rickety numbers: Volatility of university rankings and policy implications. *Research Policy*, 40, 165–177.
- Spiegel Online (2014). Deutsche Unis im "THE"-Ranking: Das Wunder von Tübingen. 02.10.2014. <http://www.spiegel.de/unispiegel/studium/uni-ranking-hochschulenim-the-ranking-a-994684.html>
- Times Higher Education. (2014). World University Rankings 2014-2015 methodology <http://www.timeshighereducation.co.uk/world-university-rankings/2014-15/worldranking/methodology>
- Tofallis, C. (2012). A different approach to university rankings. *Higher Education* 63, 1–18.
- van Raan, T. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62 (1), 133–143.
- van Raan, T., Leeuwen, T., & Visser, M. (2011). Severe language effect in university rankings: particularly Germany and France are wronged in citation-based rankings. *Scientometrics*, 88, 495–498.
- Waltman, L., Calero-Medina, C., Kosten, J., Noyons, N.C., Tijssen, R.J., & Wouters, P. (2012). The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation. CWTS Working Paper Series. <http://arxiv.org/abs/1202.3941>

The Vicious Circle of Evaluation Transparency – An Ignition Paper

Miloš Jovanović¹

¹*milos.jovanovic@int.fraunhofer.de*

Fraunhofer Institute for Technological Trend Analysis, Appelsgarten 2, 53879 Euskirchen (Germany)

Introduction

The present paper introduces a model, which describes different phases that typically occur in situations, in which a researching subject (e. g. an author, an institution, a country etc.) needs to be evaluated and in which some kind of reward (e. g. monetary in the form of a bonus or funding) is based on this evaluation. This model, the present author calls it the “vicious circle of evaluation transparency”, will be underlined by giving examples for each of its phases. In order to be able to observe a process that is described by this model, there first needs to be something that is to be evaluated, for example a research group at a university. Such a need normally comes up, when money is to be divided among different groups or focused on one. The problem of evaluation and rewarding is at the core of the model (see Figure 1).



Figure 1. The “vicious circle of evaluation transparency”-model.

Phase I – Evaluation and rewarding by subjective and intransparent criteria

The first question that might come up in such a situation is the question of how to evaluate a research group. In hierarchically organized universities the leader of a department will decide whether or not and how this group is evaluated. Very often, this person is also the one that conducts the evaluation and, based on this, determines the type and amount of a reward or funding (or some kind of penalty, if the evaluation is negative). In today’s world of vast amounts of digital data, it

might be hard for only one person to do such an evaluation. Naturally, having one person alone evaluate a group’s performance and decide on rewards will lead to a number of persons feeling unfairly evaluated, because the evaluator might not know about their achievements or their work in detail. This criticism might be alleviated in part by expanding the number of evaluators, for example by having a board of evaluators. Another possibility is to improve the transparency of the evaluation by documenting and publishing certain evaluation criteria by which the evaluated subjects can read about the evaluations and try to strive to get a better evaluation. These evaluation criteria are a first step towards phase II of the model.

Phase II – Introduction of “objective” and transparent criteria

These evaluation criteria might be subjective. For example “Quality of work” can be a criterion that is evaluated differently by different people. In order to make evaluation criteria comparable and independent of the evaluating person, “objective” criteria are often introduced. The reason why the word is put into quotation marks is due to the fact that very often these “objective” criteria are not objective at all. The introduction of “objective” and transparent criteria is a simplification of reality, an attempt to put parts of reality into some kind of a score in order to compare them with each other. Bibliometric indicators are one example of such a simplification. In many countries, different kinds of “objective” and subjective evaluation criteria have been introduced, for example in Italy (Abbott, 2009). Normally, these “objective” evaluation criteria (often in the form of different kinds of indicators) are communicated transparently. And while transparency is an important factor for these evaluations, it also leads to one problem in this phase: the fact that the evaluated subjects, in our example researchers at universities, react to the evaluation by starting to change their behavior, in order to maximize their scores in the evaluation. Of course, one reason behind evaluation is to positively influence the behavior of the evaluated researchers. But in Germany, for example, this has led to authors aiming to publish more in internationally known journals that have a US publisher and which are more general in their scope (Michels & Schmoch, 2013). This underlines the fact that authors do not base the decision in which journal they wish to publish in on scientific reasons

alone and constitutes a negative change of behavior. Also, some of the evaluated subjects might complain that the evaluation criteria do not reflect their work adequately and need to be refined. This leads to the next phase.

Phase III – Adaptation and enrichment of “objective” criteria

The need to fairly represent and evaluate researchers' work in the evaluation criteria and to adapt these in order to not allure unwanted change of behaviour leads to reforms in the evaluation system, e.g. new or a mix of indicators are proposed. The current discussion on alternative metrics is an example for phase III (e.g. in Haustein et al., 2014). The problem here is, that phase III is actually reintroducing parts of the simplification of reality, which was conducted in phase II. The evaluation criteria become more complicated again. A country example for this phase is the Czech Republic, which introduced performance-based research funding (phase II). A study by Vanacek (2014) found that the number of publications increased very quickly. He shows that in comparison to the quickly growing number of publications the quality seems to have stagnated and recommends reworking the procedure of evaluation and performance-based funding in order to increase not only the number of publications but also their quality (phase III). But for some research communities, the adaptation and enrichment of the “objective” criteria is no option. Instead, these criteria are rejected. For example, there is an ongoing discussion in the mathematical community. Authors note that bibliometric data lose “crucial information that is essential for the assessment of research”. It is pointed out that bibliometric indicators can be manipulated and lead to undesirable publishing practices (Adler, Ewing, & Taylor, 2009). The authors also dismiss reputation, as determined by surveys as a possible way of measuring the quality of a journal. The evaluation of journal editorial processes is not seen as a good way of ranking journals either. Instead, the authors recommend an “honest, careful rating of journals based on the judgment of expert mathematicians”, which is the point, where phase IV starts.

Phase IV – Removal of “objective” criteria and return to phase I

Concretely, the IMU recommends that a rating committee of 16-24 experienced and respected mathematicians should be appointed. Without going into too much detail, this committee (via various panels) is then supposed to rate the different journals and assign them to tiers (ranging from tier 1 = high quality journal to tier 4 = low-class journal) (Journal Working Group, 2011). This system is similar to the peer review process. Introducing evaluation by a committee of experts,

either by rejecting “objective” evaluation criteria or because the evaluation system has become too complicated, brings the model full circle. The evaluation has reached phase I again. One should note that in phase II of this new cycle, the criteria probably will not be the same as in the first cycle. Newly developed and more sophisticated criteria will take their place.

Conclusion

It is this author's personal opinion that the above described model of evaluation transparency not only describes a typical process in which bibliometric indicators are involved but rather evaluation processes in general. If this is the case, one may discuss possibilities to change this, since a cycle like this is not an optimal solution. An option might be the introduction of diametrically opposed evaluation criteria so that an evaluated subject could not be good in all criteria. Another idea that might serve to fan the discussion on this topic would be the introduction of a changing system of criteria, akin to the disciplines at Olympic Games. The criteria could be published a year before the evaluation takes place and would change each year. This would be a transparent system, while the evaluated researchers would not need to change their behavior in a negative way because the next year the criteria would be different. Whatever changes might be introduced, it is this author's opinion that the vicious circle has to be stopped and replaced by a different system that leads to the desired goal: a fair evaluation of research.

References

- Abbott, A. (2009). Italy introduces performance-related funding. *Nature*, 460, 559.
- Adler, R., Ewing, J. & Taylor, P. (2009). Citation Statistics: A Report from the IMU in Cooperation with the ICIAM and the IMS. *Statistical Science*, 1-14.
- Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics*, 101, 1145-1163.
- Journal Working Group (2011). *Report of the IMU/ICIAM Working Group on Journal Ranking*. Retrieved March 10, 2015 from: http://www.mathunion.org/fileadmin/IMU/Report/WG_JRP_Report_01.pdf
- Michels, C. & Schmoch, U. (2013). Impact of bibliometric studies on the publication behaviour of authors. *Scientometrics*, 98, 369-385.
- Vanacek, J. (2014). The effect of performance-based research funding on output of R&D results in the Czech Republic. *Scientometrics*, 98, 657-681.

Influence of the Research-oriented President's Competency on Research Performance in University of China -Based on the Results of Empirical Research

Li Gu¹, Liqiang Ren¹, Kun Ding¹, Wei Hu²

¹ *guli@dlut.edu.cn*

WISE Lab, School of Public Administration and Law, Dalian University of Technology, Dalian, 116024 (China)

² *huwei@moe.edu.cn*

Dept. of Personnel, Ministry of Education of China, No.37 Damucang Hutong, Xidan, Beijing, 100816 (China)

Introduction

With the gradual promotion and implementation of China's national innovation-oriented strategy, research universities are playing an irreplaceable role in leading scientific development and technological innovation. Scientific research is one of the basic functions of a research university, which cultivates high-quality innovations and supports research universities in serving their societies (Rhoads, 2014). While high-level research universities need presidents with outstanding quality and ability. Research-oriented presidents, as the scientific research managers and experts, play a very important role in constructing and developing their universities, and they also focus on talent cultivation to realize social missions.

Therefore, the research on the influence of the research-oriented president's competency on research performance has profound connotations and value, which can provide references to guide and explore the systems for selecting, cultivating and assessing research-oriented university presidents.

Method

Research-oriented presidents, as senior managers of research universities, are responsible for teaching university management and for the direct leadership of scientific research. This special position determines the universality and complexity of the factors related to empirical studies on competence characteristics (Angeles, 2014; Sydney & Frances, 2013; Liu & Xu, 2013; Snyder, 2012).

Based on the theoretical analysis of competence characteristics and in combination with the vocational characteristics and main responsibilities of research-oriented presidents, we first constructed a theoretical framework of research-oriented presidents' competence characteristics (Figure 1). Then, we designed a questionnaire system to collect data and data were analyzed using SMRT PLS2.0 software (one of the leading software tools for partial least squares structural equation modeling). The verification results show that the scale's

convergent validity was high, and it also had good discriminant validity. Finally, we used the R^2 statistic to analyze the structural model and received good explanation.

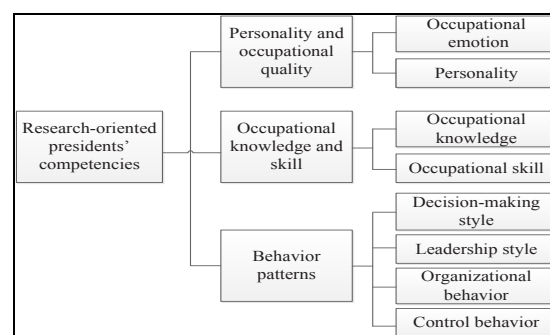


Figure 1. Theoretical Framework of Research-oriented Presidents' Competence Characteristics.

Data

This study selected research-oriented presidents of research universities as its subjects. Therefore, thirty-nine of 985 universities under China's Ministry of Education were selected for the study, and to ensure the comprehensiveness of our investigation, the selected samples included research-oriented presidents, middle management, scientific research management, professors, associate professors, lecturers, assistants, and other research personnel. The descriptive statistics (Table 1) on the study subjects were obtained via statistical data analysis.

Results

Through statistically analysing the sample data, the influence of occupational emotion, personality, occupational knowledge, occupational skill, decision-making style, leadership style, organizational behaviour and control behavior on scientific research performance was respectively checked. The results indicate that the performance had good validity. However, if organizational characteristics are used as an intervening variable, the competence characteristics of research-oriented

presidents have significant positive influences on scientific research performance.

Table 1. Descriptive Statistics on Respondents.

Measurement items		Sample size (N)	Proportion (%)
Gender	Male	292	70.4
	Female	123	29.6
Age	30 and below	37	8.9
	31–35	132	31.8
	36–40	93	22.4
	41–45	63	15.2
	46–50	41	9.9
	51–55	31	7.5
Education	56 and above	18	4.3
	College	3	0.7
	Bachelor's	31	7.5
	Master's	103	24.8
	Doctorate	276	66.5
	Others	2	0.5
Title	Assistant	98	23.9
	Lecture	92	21.9
	Associate Prof.	15	3.6
	Full Prof.	210	50.6
	Academician	0	0
	Others	0	0

Conclusion

Based on the above research results, we constructed a model of research-oriented university presidents' competence characteristics, shown in Figure 2.

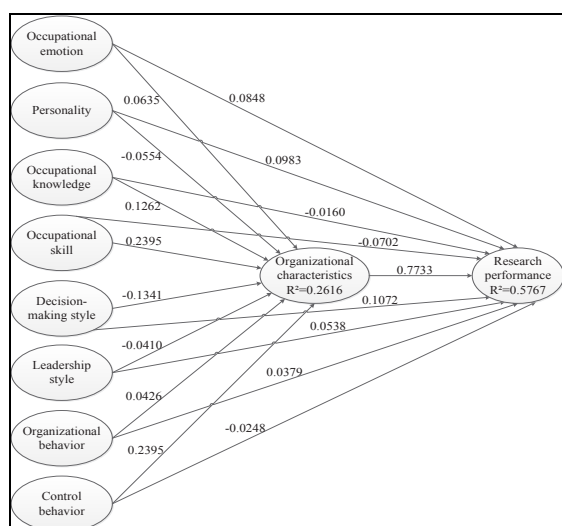


Figure 2. Relational Model of Research-oriented Presidents' Competence Characteristics and Their Universities' Research Performance.

The following conclusions can be drawn by analysing the model of research-oriented presidents' competence characteristics:

(1) From the direct effect perspective: 1) research-oriented presidents' professional emotion, personality traits, decision-making and leadership styles and organizational behavior have significant positive influences on scientific research performance. 2) Presidents' professional knowledge, professional skills and control behavior have significant negative influences on research performance, but further inspection of the analysis results reveals that the negative influence is not absolute.

(2) From the mediating effect perspective, professional emotion, professional skills, organizational behavior and control behavior have significant positive influences on organizational characteristics, whereas personality traits, professional knowledge, and decision-making and leadership styles have significant negative influences on organizational characteristics. However, organizational characteristics as intervening variables between research-oriented presidents' competence characteristics and their universities' scientific research performance can maximize the effects of the presidents' competence characteristics and have significant positive influence on research performance.

Acknowledgments

This work was supported by the project of "Specialized Research Fund for the Doctoral Program of Higher Education" (20130041120049), the project of "Fundamental Research Funds for the Central Universities" (DUT13RW409) and the project of "the Soft Science Research Project of State Intellectual Property Office – Study on Competency and Promotion Policies of Patent Attorneys from the Angle of Patent Application Quality".

References

- Angeles, M.M. (2014). Learning for a Sustainable Economy: Teaching of Green Competencies in the University. *Sustainability*, 6, 2974-2992.
- Liu, X.J. & Xu, F. (2013). A Study on the Competency Model for Industry-based Characteristic University Presidents. *Journal of National Academy of Education Administration*, 10, 60-64.
- Rhoads, R.A., Shi, X.G. & Wang, X.Y. (2014). Reform of China's Research Universities: A New Era of Global Ambition. *Education and Society*, 32, 5-28.
- Sydney, F.J. & Frances, K.K. (2013). University Presidents' Perspectives of the Knowledge and Competencies Needed in 21st Century Higher Education Leadership. *Journal of Educational Leadership in Action*, 1, 1-2.
- Snyder. (2012). Higher Education Leadership Competencies: Quantitatively Refining a Qualitative Mode. *American psychologist*, 27, 5-8.

Medical Literature Imprinting by Pharma Ghost Writing: A Scientometric Evaluation

Philippe Gorry¹

¹ philippe.gorry@u-bordeaux.fr

GREThA UMR CNRS 5113, University of Bordeaux, Av. Leon Duguit, 33608, Pessac (France)

Introduction

Misappropriation of authorship, honorary or ghost authorship, undermines academic publishing with a substantial proportion of peer-reviewed medical journals targeted (Flanagin, 1998). Pharmaceutical companies pay professional writers or medical communication companies to produce papers whilst paying other scientists or physicians to attach their names to these papers before they are published in medical or scientific journals. This ghost management is meant to support the marketing of drug products (Sismondo, 2007). Companies use this strategy to communicate competitive message, promote unproven off-label uses, and mitigate perceived drug risks (Fugh-Berman, 2010). Publication planning strategy with fraudulent practices were revealed through internal company communications in the course of the well-known Neurontin® litigation case (Vedula, 2012). Even though ghostwriting realized by pharmaceutical companies has been reported, it remains necessary to measure to what extent ghostwritten articles have impacted medical literature. Healy and Catell (2003) started to answer this question with a sample of 16 ghostwritten articles about a peculiar antidepressant. This pioneering analysis should be extended to a larger collection of ghostwritten articles as well as studied for a longer period of time.

Method

Pharma ghostwriting has been documented initially through 3 original papers: first, D. Healy and D. Cattell reported 16 ghostwritten articles in 2003, later on, A.J. Fugh-Berman (2010) reported 23 new cases, finally in 2012, Vedula and colleagues identified 13 more ghost written publications. Based on legal documents, from US district court following class action and lawsuit against pharmaceutical companies concerning several molecules: estrogen (Prempo®/Premarin®, Wyet), sertraline (Zoloft®, Pfizer), gabapentin (Neurotin®, Pfizer), and paroxetine (Paxil®, GSK), 40 more ghostwritten publications were identified. Therefore, a corpus of 92 publications were retrieved from Pubmed, Scopus or Web of Science databases, and subsequently analyzed for main bibliometric indicators. Descriptive statistics were done using Excel.

Result

A corpus of 92 ghostwritten articles was assembled, covering a period between 1997 and 2008. Two third of these cases were published between 1998 and 2000. 79 different authors have been identified. While the vast majority of them were co-author of only one ghostwriting paper, 10 authors published two ghost papers and one signed three ghost papers (data shown on the poster). 82% of the identified authors were US academics. However, authors of 10 different countries were identified as representing the main drug pharma market with the noticeable exception of Germany and Japan. Among the different affiliation of the authors, only one pharmaceutical company was identified. Most of the institutions were university with affiliated medical school (data shown on the poster).

Ghostwritten articles were published by average productive author (h-index at the time of ghost publication date: mean=15.84), with some exceptions: Bondareff W, University of Southern California, (h-index=92), Seddon JM, Tufts Medical Center, (h-index=53), Freedman MA, Medical College of Georgia & Jermain DM, Pfizer (h-index= 2). Along the 10 years observation period, there is no noticeable variation in the productivity of the authors (data shown on the poster). Indeed average author h-index reach 29.13 in year 2013.

The corpus covers a large spectrum of medical specialties. However, it is interesting to point out that more than a third of ghostwritten papers concern psychiatry and mental illness (Figure 1).

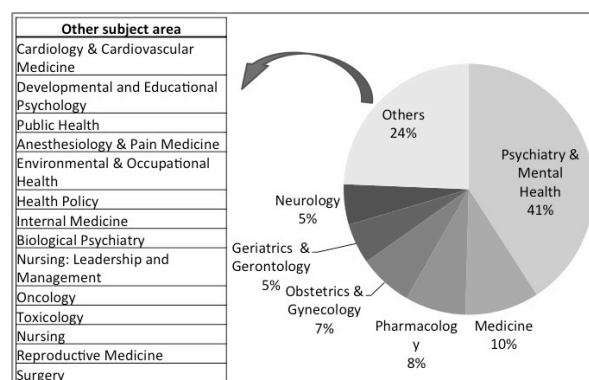


Figure 1. Distribution of ghost written articles by medical specialties.

Publication of ghost articles were scattered throughout 51 different journals. Among these source titles, there are four psychiatric journals, with various impact factor (IF), accounting for a third of the ghostwritten articles (Figure 2 and Table 1).

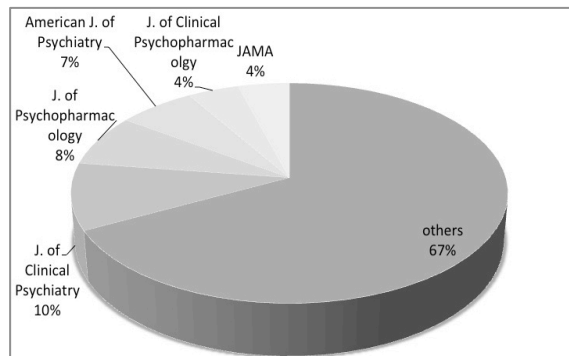


Figure 2. Distribution of ghost written articles by journals.

Table 1. List of the main journal publishing ghost written articles with their impact factor.

Journal	Ghost pub.	SJR impact factor at publication date
Journal of Clinical Psychiatry	9	1,787307692
Journal of Psychopharmacology	7	1,142571429
American Journal of Psychiatry	6	3,599
Journal of Clinical Psychopharmacology	4	1,6045
Journal of the American Medical Association	4	3,82875

The average IF of journals where ghostwritten articles are published is in the low-medium range (mean IF=2.51, median IF=1.81). Sometime, there are published in very low IF journal (ex: Climacteric IF=0.091).

Finally, the last evaluation concerns the number of year during which a ghostwritten article can be cited since the date of publication. (Figure 3; no ghostwritten article have been published in 2007). Year after year, ghostwritten articles have on an average 84% chance to be cited.

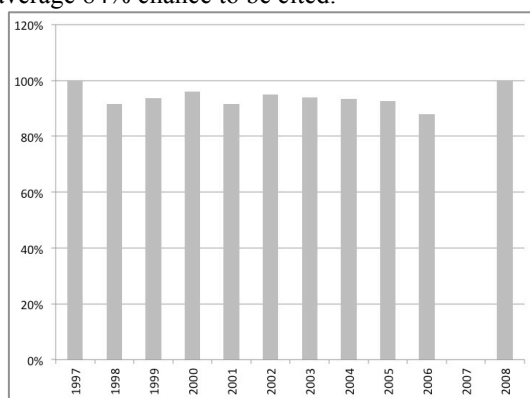


Figure 3. Probability of a ghost written articles to be cited once year since the publication.

On long range, the average ghostwritten article IF is much higher than the average journal IF. Indeed a ghostwritten article is about 10 times more cited

than any article published in the same journal (Table 2).

Table 2. Statistics difference between ghost written & journal article impact factors.

	Ghost Article impact factor	Journal impact factor
mean	7,24	2,68
max	68,13	8,73
min	0,31	0,09

Discussion

With this study, we have been able to conduct a bibliometric analysis on a large number of ghost articles, over a long period of time. Overall, ghostwritten articles are published by average productive author, in low IF journals; they are cited during a long period of time and therefore have a high number of citations (Table 3). Thus, ghostwritten articles might influence the medical community and its practice, which subsequently raises public health concerns.

Table 3. Main bibliometric indicators of ghost written articles.

	Journal impact factor	Ghost written article citations	Last year citation	Total number year citations	author h index
Mean	2,697	84,951	2013	13	29,731
Max	6,984	351	2014	16	68
Min	0,091	4	2008	8	4

Despite numerous declarations by medical journal editors and the conduct of ethics declared by professional medical writers, we would like to underline that none of these ghostwritten articles involved in lawsuit case have been retracted whilst companies have been sentenced by Justice.

Moreover the efficiency of ghostwriting publication strategy could be questioned since only a third of articles have an impact superior to what would be expected. Therefore the return on investment for the pharmaceutical industry might be very low, especially regarding the risk of litigation and the disclosure of such fraudulent marketing practices.

References

- Flanagin, A., *et al.* (1998). Prevalence of articles with honorary authors and ghost authors in peer-reviewed medical journals. *Journal of the American Medical Association*, 280, 222-224.
- Fugh-Berman, A.J. (2010). The Haunting of Medical Journals: How Ghostwriting Sold "HRT" *PLoS Medicine*, 7, e1000335.
- Healy, D. & Cattell, D. (2003). Interface between authorship, industry and science in the domain of therapeutics. *British J. of Psychiatry*, 183, 22-27.
- Sismondo, S. (2007). Ghost management: How much of the medical literature is shaped behind the scenes by the pharmaceutical industry? *PLoS Medicine*, 4, 1429-1433.
- Vedula, S.S. *et al.*, (2012). Implementation of a publication strategy in the context of reporting biases. A case study based on new documents from Neurontin litigation. *Trials*, 13, 136.

Are scientists really publishing more?

Daniele Fanelli¹, and Vincent Larivière²

¹ dfanelli@stanford.edu

METRICS - Meta-Research Innovation Center at Stanford, Stanford University, 1070 Arastradero Road, Palo Alto, CA 94304

² vincent.lariviere@umontreal.ca

Université de Montréal, École de bibliothéconomie et des sciences de l'information, C.P. 6128, Succ. Centre-Ville, H3C 3J7 Montreal, Qc. (Canada)

Université du Québec à Montréal, Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Observatoire des Sciences et des Technologies (OST), CP 8888, Succ. Centre-Ville, H3C 3P8 Montreal, Qc. (Canada)

Introduction

The success of researchers and research institutions is increasingly determined by measurable aspects of their performance, in particular the quantity and citation-impact of their publications. The effects that these growing “pressures to publish” might have on publication and research practices are a matter of growing concern and increasing academic interest (de Winter & Dodou, 2014; Fanelli, 2010, 2012, 2013; Tijdink, Vergouwen, & Smulders, 2013; van Dalen & Henkens, 2012).

Much criticisms and concern has been expressed, in particular, for the risk of overemphasising the quantity of a scientist's publication record at the expense of its quality. In order to show a longer lists of publications in their CVs, it is commonly hypothesised, scientists might increasingly resort to questionable practices such as inappropriately subdividing (“salami slicing”) their results, publishing trivial and incomplete studies, conducting research hastily and sloppily, selecting out of their findings those that are least “publishable”, or even resorting to outright scientific misconduct in the form of duplicate publication, plagiarism and data fabrication (e.g. Angell, 1986; Hayer et al., 2013).

Performance-evaluation policies of institutions in various countries have responded to these concerns by formally removing any quantitative consideration from their performance assessments (e.g. VSNU, 2015). However, there is little evidence to support these policies. No study, in particular, has ever verified whether scientists are have actually responded to growing pressures by churning out more papers. We present preliminary results of a project aimed at filling this gap in the literature.

Methods

We identified individual researchers who published in the Web of Science across the 20th century by selecting all authors identified by three initials (first name and two middle names, plus surname, e.g.

Vleminckx-SGE), which reduces the likelihood that these researchers have homonyms. From this initial sample we selected authors who had at least two publications, and from these we then selected authors whose publications spanned a period of at least 15 years. For each of these authors we then counted the total number of papers published in the first 15 years of activity – the period where pressures to publish are hypothesised to be stronger – and we also measured the average number of co-authors.

Results

The raw number of papers published by individual authors has grown very rapidly across the century (Fig. 1). Fractional productivity, however, as measured by dividing the author's total number of papers by the average number of co-authors, shows a net decline (Fig. 2).

Discussion

Although still preliminary, these results suggest that our beliefs about the effects of pressures to publish might be partially incorrect. Authors might have responded to growing performance expectations not, as commonly believed, by subdividing or trivializing their results or by multiplying their effort at the expense of other activities, but by enlarging their network of collaborations in order to make ever smaller contributions to a growing number of projects. Since neither publication nor citation metrics are counted fractionally, this strategy allows scientist to increase their measurable publication rate without necessarily increasing their total research effort.

If scientists' net effort devoted to research is not increasing, then concerns for growing “salami slicing” and other questionable practices might be unjustified. Explanations for recent evidence that retraction and correction rates are growing (Fang & Casadevall, 2011), that publication bias is growing (Fanelli, 2012) and that research bias might be higher in scientifically productive countries (Fanelli, 2010) might need revising. And policies that are currently de-emphasizing “quantity” in

favour of “quality” (e.g. VSNU, 2015) might not have a solid basis in evidence, and could therefore be ineffective or even damaging.

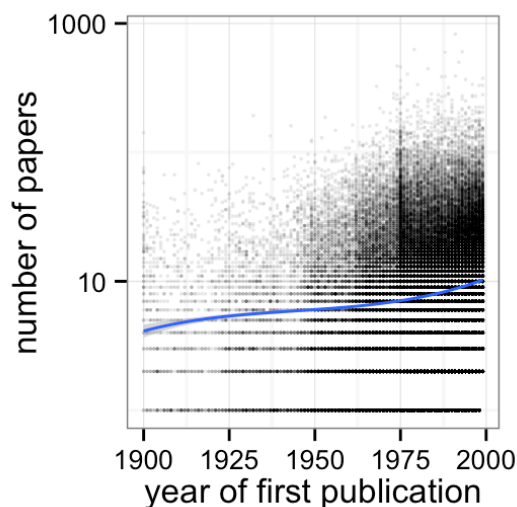


Figure 1. Total number of papers published during the first 15 years of career (N= 70,310). Blue line: cubic polynomial regression fit, with grey areas representing 95%CI.

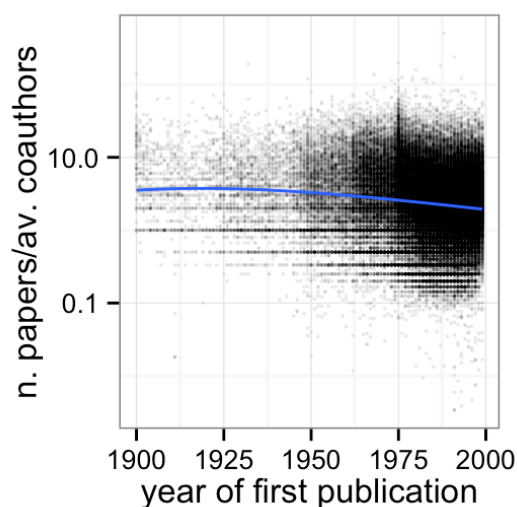


Figure 2. Ratio of total number of papers to average number of co-authors during the first 15 years of career (N= 70,310). Blue line: cubic polynomial regression fit, with grey areas representing 95%CI.

Several limitations to these results, however, remain to be addressed. First, since the likelihood of having two middle names is very unequally distributed amongst countries, our sample might not be sufficiently representative of the corpus of literature in the Web of Science. Second, our method might not be sufficiently robust against disambiguation errors for names from South-East Asian countries, a problem which might have

skewed our results. Third, the Web of Science database does not cover a significant proportion of the literature, and its coverage varies by discipline and across the years. Future work will aim at adjusting for these factors, in order to verify whether scientists are actually publishing more or just collaborating more extensively.

Acknowledgments

The authors acknowledge funding from the Canada Research Chairs program as well as from the Social Sciences and Humanities Research Council of Canada.

References

- Angell, M. (1986). Publish or Perish - A proposal. *Annals of Internal Medicine*, 104(2), 261-262.
- de Winter, J., & Dodou, D. (2014). A surge of p-values between 0.040 and 0.049 in recent decades (but negative results are increasing rapidly too). *PeerJ, PrePrints* (2), e447v443
- Fanelli, D. (2010). Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data. *Plos One*, 5(4). doi: 10.1371/journal.pone.0010271
- Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics*, 90(3), 891-904. doi: DOI 10.1007/s11192-011-0494-7
- Fanelli, D. (2013). Why Growing Retractions Are (Mostly) a Good Sign. *PLoS Med*, 10(12), e1001563. doi: 10.1371/journal.pmed.1001563
- Fang, F. C., & Casadevall, A. (2011). Retracted Science and the Retraction Index. *Infection and Immunity*, 79(10), 3855-3859. doi: 10.1128/iai.05661-11
- Hayer, C.-A., Kaemingk, M., Breeggemann, J. J., Dembkowski, D., Deslauriers, D., & Rapp, T. (2013). Pressures to Publish: Catalysts for the Loss of Scientific Writing Integrity? *Fisheries*, 38(8), 352-355. doi: 10.1080/03632415.2013.813845
- Tijdkink, J. K., Vergouwen, A. C. M., & Smulders, Y. M. (2013). Publication Pressure and Burn Out among Dutch Medical Professors: A Nationwide Survey. *PLoS ONE*, 8(9), 6. doi: 10.1371/journal.pone.0073381
- van Dalen, H. P., & Henkens, K. (2012). Intended and Unintended Consequences of a Publish-or-Perish Culture: A Worldwide Survey. *Journal of the American Society for Information Science and Technology*, 63(7), 1282-1293. doi: 10.1002/asi.22636
- VSNU (2015). Protocol for Research Assessments in the Netherlands (2015).



PATENT ANALYSIS

COUNTRY LEVEL STUDIES

Tapping into Scientific Knowledge Flows via Semantic Links

Saeed-UI Hassan¹ and Peter Haddawy²

¹*saeed-ul-hassan@itu.edu.pk*

Information Technology University, 346-B, Ferozepur Road, Lahore (Pakistan)

²*peter.had@mahidol.ac.th*

Faculty of ICT, Mahidol University, 999 Phuttamonthon 4 Rd, Salaya, Nakhonpathom 73170 (Thailand)

Abstract

We present a new technique to semantically analyze knowledge flows between countries by using bibliometric data. Using a new approach to keyword-based clustering, the technique identifies the main topics of the research output of a country, as well as the main topics of the citing research of other countries. In this way it provides insight into how research produced by one country is used by others. We present a case study to illustrate the use of our proposed technique in the subject area of Renewable Energy during 2005-2010 using data from the Scopus database. We compare the Japanese and Chinese papers that cite the scientific literature produced by researchers from the United States in order to show the difference in the use of same knowledge. While the Japanese researchers focus on research areas such as efficient use of Photovoltaics and Superconductors, Chinese researchers focus in areas related to Power Systems, Power Management and Hydrogen Production. Such analyses may be helpful in establishing more effective multi-national research collaboration.

Conference Topics

Methods and techniques; Country-level studies

Introduction

The research collaboration facilitated by the Internet and the greatly increased global mobility of researchers have resulted in a new highly dynamic global marketplace for ideas. The possession of knowledge, the value of which depreciates at an increasingly rapid rate, is no longer as valuable as the ability to participate in the knowledge flows associated with these marketplaces. As observed by Hagel et al. (2009) in the context of business competitiveness, “Knowledge flows – which occur in any social, fluid environment where learning and collaboration can take place – are quickly becoming one of the most crucial sources of value creation”. Similarly in Science, understanding a research landscape increasingly requires understanding the dynamics of the relevant knowledge flows.

International scientific leadership and influence are commonly viewed as important measures of a country’s scientific intellectual strength. This has traditionally been measured in terms of international scientific collaboration and the ability of a country to attract strong researchers and graduate students from abroad. But a further, more direct measure is the extent to which results generated by a country’s researchers are influencing and being utilized by researchers abroad, particularly researchers who are not yet directly collaborating with that country’s researchers.

In this paper we present a new technique to measure and semantically analyze knowledge flows between countries by using publication and citation data. We select a set of papers authored by the researches of a given source country. Further, we identify the papers cited by the papers only authored by researchers from outside the source country. We cluster these internationally cited papers to identify the main topics. Then, we procure the sets of papers (authored by researchers outside the given country) citing each of the topic clusters. Finally, we in turn cluster each set of citing papers to again identify main topics in order to identify how the knowledge from the topics in the cited papers is being used.

Related Work

In bibliometrics there have been efforts to measure knowledge flows using scientific literature at different levels of detail, namely: among scientists, among journals, among subject categories, among institutions and among countries.

Zhuge (2006) argues that ideas in a scientific article inspire new ideas, which will be recorded and published as new articles after peer review. Therefore, citations between scientific articles imply a knowledge flow from the authors of the article being cited to the authors of the articles that cite it. Zhou and Leydesdorff (2007) use journal-journal citation analysis to investigate international visibility of journals. Zhou et al. (2010) also use journal-journal citation analysis to study the specialization of a research community within a discipline. Johannes and Guenter (2001) measure knowledge export and international visibility of journals by determining the unique subject fields to which the citing journals have been assigned and the unique countries to which the citing authors belong, respectively.

Rowlands (2002) proposes a method to measure the spread of scientific knowledge that is published in a journal. He focuses on journals as units of spread and introduces an indicator to measure the spread of knowledge by looking at the number of different journals that cite the papers published in the primary journal, as shown in Equation 1.

$$RDI = \frac{U}{Cit}, \quad (1)$$

where U stands for the number journals that cite the papers published in the primary journal in a given time window (say T). Cit is the total number of citations received by the articles in the primary journal in T time window and the notion RDI is for Rowlands Diffusion Index. Naturally, diffusion can only increase in an absolute sense, however, empirical results show that the diffusion index proposed by Rowlands is negatively correlated with the total number of citations received (Rowlands, 2002). This leads Frandsen (2004) to provide a different diffusion index, as shown in Equation 2.

$$FDI = \frac{U}{Pub}, \quad (2)$$

where Pub stands for total number of publications in the primary journal, U is the same as above and FDI stands for Frandsen Diffusion Index. Note that Cit is replaced by Pub (i.e. publications). When publications do not change, the Frandsen Diffusion Index cannot decrease, and thus, the Frandsen Diffusion Index is positively correlated with the total number of citations.

Burrell (1991, 1992, 2005 and 2006) shows that the Leimkuhler Curve can provide an intuitive visual representation for the Gini Coefficient Index in giving graphical and numerical summaries of the concentration of bibliometric distributions. Guan and Ma (2007) illustrate the use of the Leimkuhler Curve to reveal the impacts of research outputs of countries. Using the Gini index, Liu and Rousseau (2010) study knowledge diffusion through publications and citations, as shown in Equation 3.

$$G = \frac{2q - 1}{N}, \text{ where}$$

$$q = \sum_{i=1}^N i \frac{x_i}{M}$$

$$M = \sum_{i=1}^N x_i \quad (3)$$

N denotes the number of subject categories, and x_i denotes number of citations in journals mapped with a given subject category i . Note that the Gini index (Burrell, 1992, 2005) can be equally computed using Equation 4.

$$G = 1 - \frac{\sum_{j=1}^{\infty} (r(j))^2}{N \cdot M}, \quad (4)$$

where M and N are the same as in Equation 3, $r(j)$ stands for the number of subject areas with at least j citations and the sum is finite as there is always a subject category with the largest number of citations. Note that Gini based indexes can only characterize the knowledge diffusion and do not quantify the volume of knowledge flow.

Ingwersen et al. (2000) present international citations as an indicator to measure export of knowledge produced by institutions. They measure knowledge export of institutes by calculating the proportion of citations received by a given institute from other countries (outside the host country where the institute is located) relative to total citations received by the institute. Using citation exchange among the scientific articles, we introduce a notion of International Scholarly Impact of Scientific Research (ISISR) to measure international knowledge flows among countries and institutions (Hassan & Haddawy, 2013). However, the measure of ISISR only quantifies knowledge flows and does not elucidate the contents of knowledge that flows across the countries.

The above survey discusses the salient research to quantitatively measure knowledge flows using bibliometric data. However, we believe that apart from the quantitative measures it is extremely important to analyze the contents of the knowledge flows. The scientific work of Zhuge (2009, 2010, 2011 & 2012) sets the theoretical base of semantic analysis in order to extract knowledge from large scale corpus.

Methodology

This section presents analytical techniques used to semantically analyze the knowledge flow from a given source country. We consider a set of papers P authored by the researchers of a given source country in a given subject area in a given time window. Among the selected papers, we identify the papers P cited by the papers only authored by researchers from outside the source country. We cluster the papers from P to identify the main topics. We procure the sets of papers (authored by researchers outside the given country) citing each of the topic clusters. Next, we in turn cluster each set of citing papers to again identify main topics in order to identify how the knowledge from the topics in the cited papers is being used. The research topics are identified using our proposed Topic with Distance Matrix (TDM) model, an extension of the Latent Dirichlet Allocation (LDA) model proposed by Blei et al (2003).

A number of approaches to model scientific paper content have been proposed (Blei et al., 2003; Hofmann, 1999). These approaches are based upon the idea that the probability distribution over words in a paper can be expressed as a mixture of topics, where each topic is a probability distribution over words. We utilize one such popular model, LDA, proposed by Blei et al. (2003). In LDA, the generation of a paper collection is modeled as a three step process. First, for a given paper, a distribution over topics is sampled from a Dirichlet distribution. Then, for each word in the paper, a single topic is selected according to this distribution. At Last, each word is sampled from a multinomial distribution over words specific to the sampled topic.

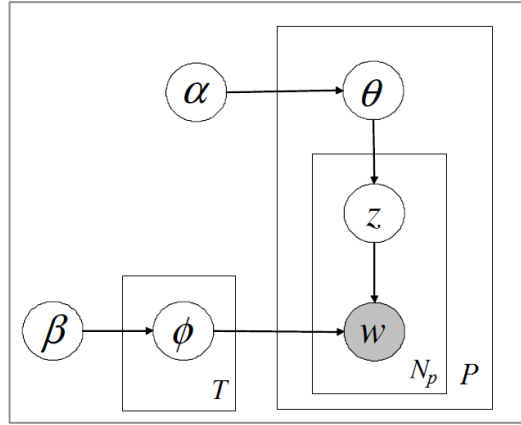


Figure 1. Latent Dirichlet Allocation (LDA) Model.

Using plate notation, the generative process corresponding to the hierarchical Bayesian model is shown in Figure 1. In this model, Φ stands for the matrix of topic distributions for each of T topics being selected independently from a symmetric Dirichlet prior (β). Θ is the matrix of paper specific mixture weights for these T topics, each being drawn independently from a symmetric Dirichlet prior (α). For each word, z denotes the topic responsible for generating that word, drawn from the Θ distribution for that paper, and w is the word itself, drawn from the topic distribution Φ corresponding to z . A paper p is a vector of N_p words, w_d , where each w_{id} is chosen from a vocabulary of size V and P is a collection of papers.

Estimating Θ and Φ provides information about the topics that participate in a publication corpus and the weights of those topics in each paper respectively. A variety of algorithms have been used to estimate these parameters, including variational inference (Blei et al., 2003), expectation propagation (Minka & Lafferty, 2002), and Gibbs sampling (Griffiths & Steyvers, 2004). To induce the probability distribution of Θ and Φ , LDA uses Gibbs Sampling which starts from randomly selected initial states and then revises distributions by changing topics to find correct distributions. Finally, the model provides topic-word relationship by the vector formed probabilistic representations.

Using the LDA, we obtain topic vectors where each value in the vector is associated with a given word that shows the probability of the word occurring under the given topic. For instance, vector T_1 (word₁: 0.3, word₂: 0.1, word₃: 0.2, ..., word_n: 0.8) shows the probability distribution of all n words for the given topic t_1 . Using this information, we represent each paper (from the set P) in the form of a vector where each value in the vector represents the probability distribution of a given word from vocabulary V in the paper for the topic under consideration (say t_1). For instance, P_1 (word₁: 0.4, word₂: 0.2, word₃: 0.0, ..., word_n: 0.7) shows the probability distribution of words in the paper p_1 for the topic t_1 . Note that if a word from V does not appear in p_1 then we assign default zero probability for that word.

Using the Minkowski distance between a given paper-vector P and topic-vector T , we choose papers in order to classify them as belonging to a specific topic (see Equation 5).

$$D = \sqrt{\sum_{i=1}^n |a_i - t_i|^2}, \quad (5)$$

where a_i denotes the probability of the term i in paper p_l for the given topic T , and t_i denotes the probability of term i for the topic T . In order to obtain a set of papers relevant to topic T , a threshold TH is applied with the given percentage of the distance between the minimum and the maximum distance of paper vectors from T . Our experimental results show that the highest F-measure is achieved with $TH = 25\%$. The size of a topic is determined by the

number of papers associated with it. The numbers of topics are determined by computing inter and intra topic similarity. We minimize inter topic similarity and maximize intra topic similarity to obtain the optimal number of topics. To compute the inter similarity between two topic, we use the Jaccard distance index (Jaccard, 1901).

Case Study: Semantic Analysis of Knowledge Flows across Countries in the Field of Renewable Energy

Dataset

We present a case study to illustrate the use of our technique in the subject area Renewable Energy. Using All Science Journal Classification (ASJC), we procured 46,518 publications (journal articles, reviews and conference papers) classified as Renewable Energy, a subarea of Energy(all) from the Scopus database during the time period 2005-2010

We procure 8,590 papers (P') (journal articles, reviews and conference papers) published by researchers from the United States. Among the selected set of papers P' , we select 4,362 papers (P) which are cited by papers authored only by researchers from other countries. Further, we select candidate terms to represent each paper. In order to procure such terms, we use author defined keywords from the selected papers. In addition, we extract noun terms from the abstracts and titles of the papers using SharpNLP (<http://www.codeplex.com/sharpnlp>). We then identify synonyms of the selected noun terms using WordNet 3.0 (<http://wordnet.princeton.edu/>) and include them as candidate terms as well. Next, we apply the Porter Stemming algorithm (<http://tartarus.org/martin/PorterStemmer/>) to stem all the selected candidate terms. Finally, we feed this data to our TDM model.

Research Topics Cited by Researchers from Outside the United States in the Field of Renewable Energy

Figure 2 shows four research topics in the field of Renewable Energy cited by researchers from outside the United States. Using Wordle.Net (<http://www.wordle.net/>), we visualize the contents in each topic. Here, each topic is represented with the most frequently occurring author defined keywords collected from the papers in a given topic. The number of papers belonging to a specific research topic and the size of each research topic are written next to its respective topic. The research topics 1 and 4 are the largest topics cited by researchers from outside the United States. The topic#1 is the largest topic, containing 44% of the 4,362 papers. This topic covers research work related to Solar Cells, Solid Oxide Fuel Cells (SOFC) and Proton Exchange Membrane Fuel Cells (PEMFC). The topic#2 is related to Hydrogen Production. This topic also covers research related to Steam Reforming, a method for producing hydrogen, carbon monoxide, or other useful products from hydrocarbon fuels such as natural gas. Finally, the topic#3 is about Li-ion batteries. Li-ion batteries are an important type of rechargeable battery, particularly used in mobile devices. Finally, the topic#4 covers research related to Sustainable Management. Next we explore how the researcher from different countries cites the knowledge produced by the United States.

Research Topics of the Publications Produced by Chinese and Japanese Researchers that Cite Papers Authored by Researchers from the United States

To understand the difference in the use of the same knowledge, we further analyse that how the scientific knowledge diffuses into other research topics used by different research communities. We compare publications of the researchers from China and Japan that cite the same knowledge produced by the researchers from the United States. We select topic#1 from Figure 2 (the largest topic cited by the researchers from outside the United States in the field of Renewable Energy during 2005-2010). This topic covers research topics related to Solar

Figure 3 shows research topics of the scientific knowledge produced by the Chinese researchers during 2005-2010 that cite topic#1 in Figure 2. In Figure 3, topic#1 mainly covers research related to Power Systems, Energy Management and Production. This topic is the largest topic which contains 53% papers out of 318. The topic#2 which contains 47% of the papers mainly focuses on Hydrogen Production.

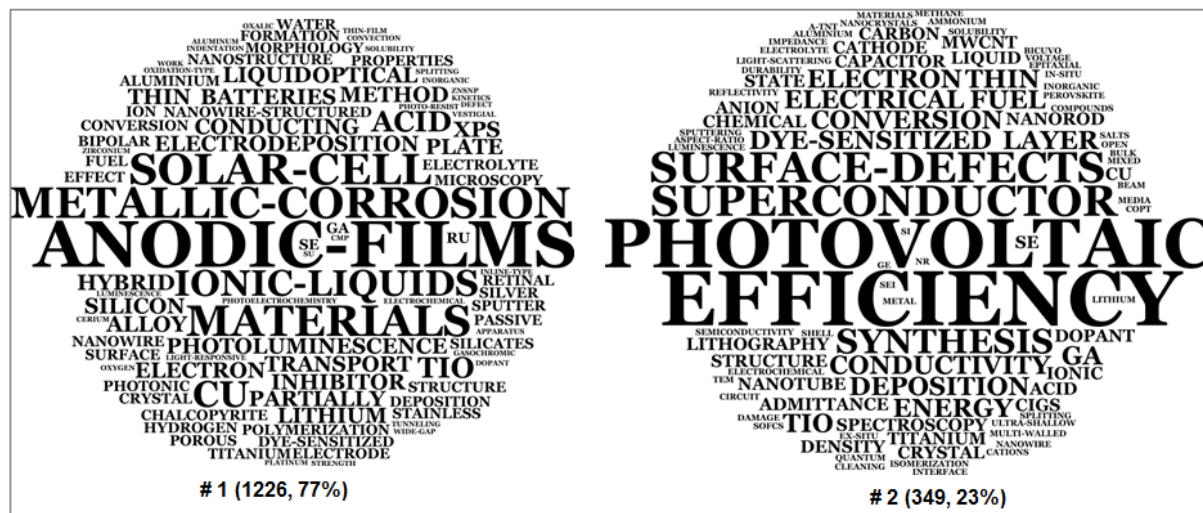


Figure 4. Research Topics of the Scientific Knowledge Produced by the Japanese Researchers (during 2005-2010) that cite topic#1 in Figure 2.

Figure 4 shows research topics of the scientific knowledge produced by the Japanese researchers during 2005-2010 that cite topic#1 in Figure 2. In contrast with China, the Japanese research community utilizes the same knowledge (produced by the United States) in rather different research themes. The Japanese researchers focus on topics related to Metallic Corrosion and Anodic Oxide Films (see topic#1 in Figure 4). Interestingly, we also find another topic (topic#2: 55 papers) describing the efficient use of Photovoltaics, Dye-sensitized Solar Cells and Superconductors. Note that Superconductors play a vital role in providing low-cost renewable energy.

Concluding Remarks

In this paper we have presented a new topic model with distance matrix, called TDM, to semantically analyze knowledge flows across countries by using publication and citation data. We have also presented a case study to illustrate the use of our proposed techniques in the subject area of Renewable Energy during 2005-2010 using data from the Scopus database. We have compared the Japanese and Chinese papers that cite the same scientific literature produced by the researchers from the United States in order to show the difference in the use of same knowledge. The study has shown that Japanese researchers focus in research areas such as efficient use of Photovoltaics, and Superconductors (to produce low-cost renewable energy). In contrast with the Japanese researchers, Chinese researchers focus in the areas of Power Systems, Power Management and Hydrogen Production.

The method of semantic analysis presented in this paper provides an understanding of the internationality of research not provided by studies of researcher mobility and co-authorship patterns. Our case study highlights the diversity in the ways that research produced by a country may be used in different international contexts, even within a relatively narrow research area. Such analyses may be helpful in establishing more effective multi-national research collaboration and in aligning collaboration with national priorities.

References

- Blei, M., Ng, A. & Jordan, M. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Burrell, Q. L. (1991). The Bradford distribution and the Gini index. *Scientometrics*, 21, 181–194.
- Burrell, Q.L. (1992). The Gini index and the Leimkuhler curve for bibliometric processes. *Information Processing and Management*, 28(1), 19-33.
- Burrell, Q.L. (2005). Measuring similarity of concentration between different informetric distributions: Two new approaches. *Journal of the American Society for Information Science and Technology*, 56(7), 704-714.
- Burrell, Q.L. (2006). Measuring concentration within and co-concentration between informetric distributions: An empirical study. *Scientometrics*, 68(3), 441-456.
- Frandsen, T. (2004) Journal diffusion factors: A measure of diffusion?. *Aslib Proceedings*, 56(1), 5-11.
- Griffiths, T. & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(1), 5228–5235.
- Guan, J. & Ma, N. (2007). A bibliometric study of China's semiconductor literature compared with some other major Asian countries, *Scientometrics*, 70(1), 107-124.
- Hagel, J., Brown, J. & Davison, L. (2009). *Measuring the forces of long-term change: The 2009 shift index*. Deloitte Development LLC.
- Hassan, S. & Haddawy, P. (2013). Measuring international knowledge flows and scholarly impact of scientific research, *Scientometrics*, 94(1), 163–179.
- Ingwersen, P., Larsen, B. & Wormell, I. (2000). Applying diachronic citation analysis to ongoing research program evaluations. In B. Cronin & H.B. Atkins (Ed.), *The Web of Knowledge* (pp. 373-387). Medford, N.J.: Information Today, Inc. & American Society for Information Science.
- Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura, *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37, 547–579.
- Johannes, S. & Guenter, G. (2001). Citation rates, knowledge export and international visibility of dermatology journals listed and not listed in the Journal Citation Reports. *Scientometrics*, 50(3), 483-502.
- Liu, Y. & Rousseau, R. (2010). Knowledge diffusion through publications and citations: A case study using eSI-fields as unit of diffusion. *Journal of the American Society for Information Science and Technology*, 61(2), 340-351.
- Minka, T., & Lafferty, J. (2002). Expectation-propagation for the generative aspect model. *Proceedings of the Eighteenth Conf. on Uncertainty in Artificial Intelligence*, 352–359.
- Rowlands, I. (2002). Journal diffusion factor: A new approach to measuring research influence. *Aslib Proceedings*, 54(2), 77–84.
- Steyvers, M., Smyth, P., Rosen-Zvi, M., & Griffiths, T. (2004). Probabilistic Author-Topic Models for Information Discovery. *The Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Seattle, Washington.
- Zhou, P. & Leydesdorff, L. (2007). A comparison between the China scientific and technical papers and citations database and the Science Citation Index in terms of journal hierarchies and inter-journal citation relations. *Journal of the American Society for Information Science and Technology*, 58(2), 223-236.
- Zhou, P., Su, X., & Leydesdorff, L. (2010). A comparative study on communication structures of Chinese journals in the social sciences. *Journal of the American Society for Information Science and Technology*, 61(7), 1360-1376.
- Zhou, P. & Leydesdorff, L. (2007). A comparison between the China scientific and technical papers and citations database and the Science Citation Index in terms of journal hierarchies and inter-journal citation relations. *Journal of the American Society for Information Science and Technology*, 58(2), 223-236.
- Zhou, P., Su, X. & Leydesdorff, L. (2010). A comparative study on communication structures of Chinese journals in the social sciences. *Journal of the American Society for Information Science and Technology*, 61(7), 1360-1376.
- Zhuge, H. (2006). Discovery of knowledge flow in science. *Communications of the ACM*, 89(5), 101-107.
- Zhuge, H. (2009). Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning, *IEEE Transactions on Knowledge and Data Engineering*, 21(6), 785-799.
- Zhuge, H. (2010). Interactive Semantics, *Artificial Intelligence*, 174, 190-204.
- Zhuge, H. (2011). Semantic linking through spaces for cyber-physical-socio intelligence: A methodology, *Artificial Intelligence*, 175, 988-1019.
- Zhuge, H. (2012). Knowledge Flow, Chapter 5 in *The Knowledge Grid - Toward Cyber-Physical Society*, 2nd Edition, *World Scientific Publishing Co.*, Singapore.

Causal Connections between Scientometric Indicators: Which Ones Best Explain High-Technology Manufacturing Outputs?

R. D. Shelton¹, T. R. Fadel², P. Foland³

¹*shelton@wtec.org*

WTEC, 1653 Lititz Pike #417, Lancaster, PA 17601 (USA)

²*tarek.r.fadel@gmail.com*

³*pfoland14@gmail.com*

ITRI, 518 S. Camp Meade Rd., Baltimore, MD 21090 (USA)

Abstract

Scientometric models can connect indicators via cross-country correlations, but these are not enough to assert causality. Sometimes a causal connection can be argued from the physical process. In other cases the causality or its direction is not clear, and the Granger test is often used to clarify the connection. Here it was shown that gross expenditures on R&D (GERD) Granger caused scientific papers in the U.S., EU, and some others, which has policy implications. Granger causality also reinforces earlier findings on why the EU passed the U.S. in papers in the mid-1990s. Downstream, it is difficult to prove the connection between research and gross domestic product (GDP), since the contributions of science are diluted by other factors. New data allows a focus on a sector that is more closely associated with science: high technology (HT) manufacturing outputs. This value-added data permits more accurate models for today's international supply chains. Correlations show that business expenditures on R&D (BERD) and scientific indicators like patents are closely connected with HT manufacturing outputs. However for BERD, either direction of causality is plausible, and enough countries had significant results to show that causality can indeed be in either direction. The connections between papers and patents with HT manufacturing were also investigated; in several countries patents could be said to have Granger caused HT manufacturing.

Conference Topic

Country-level studies

Introduction

Correlation does not imply causality, unless it can be augmented with other evidence. Many researchers have found strong cross-country correlations between national R&D funding and intermediate indicators like papers and patents. These findings bolster the policy argument that researchers deserve more funding, but may sound self-serving. Here however, there is a convincing physical argument that there is philosophical causality. Everyone knows that it takes resources to do research. In some "big science" fields like ITER and CERN, it takes international consortia to provide the necessary big funding. Even the lonely bibliometrician needs a computer, data and Internet access, time to do the work, and travel funds to present the results in some pleasant clime.

Downstream in the innovation process, many researchers have also tried to connect those papers and patents to outputs like gross domestic product (GDP), with mixed success. Here the physical connection is not so clear, because science is only one of many factors that are involved. For example, several Asian nations became export powerhouses with skyrocketing GDPs, based initially on imported technologies, which were not reflected in their national papers and patents. Instead, the "New Economic Geography" developed by Paul Krugman (1991) identifies the most significant factors for location of manufacturing, and location of R&D is not high on the list. (He won the 2008 Nobel Prize for this work.) Once prosperous, these nations did invest in indigenous innovation.

In these more difficult cases, analysts rely on statistical tests to provide some evidence of causality. The most common test was devised by Clive Granger (1969). (He also won the Nobel Prize, in 2003.) It is applied to two time series, which the analyst suspects may be related. In simplified terms, a time series x can be said to "Granger cause" a second time series y if the additional knowledge of x allows a significantly better prediction of y than simply the past history of y . The Granger test function is available in several statistical programs; the open source R software was used here (R Core Team, 2014). In the R version, the model order k is the same for both x and y . The null hypothesis that x does not Granger-cause y is not rejected, if and only if no lagged values of x are retained in the regression.

Let y and x be stationary time series. To test the null hypothesis that x does not Granger-cause y , one first finds the proper lagged values of y to include in an autoregression of y :

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_k y_{t-k} + \text{residual}_t$$

Next, the autoregression is augmented by adding lagged values of x :

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_k y_{t-k} + b_1 x_{t-1} + \dots + b_k x_{t-k} + \text{residual}_t$$

One retains in this regression all lagged values of x that are individually significant according to their t -statistics, provided that they collectively add explanatory power to the regression according to an F -test; adapted from Seth (2007). Here the smallest model order that produces significant results is preferred.

Granger testing is not a panacea. It requires that both series be stationary, and scientometric series usually fail the standard Augmented Dickey Freeman (ADF) test. This is often because they have trends such as inflation, population growth, or just more journals in the Science Citation Index (SCI). One normally has to de-trend series, usually by differencing them one or more times. Even when both series are stationary, the Granger test often fails, or worse, shows bi-directional causality, raising more questions than it answers. Furthermore, Granger causality is based on a postulate that cause must precede effect, but is this always true? In the stock market, the prospect of future events, like increased earnings, can influence present stock prices. Thus, one cannot prove true philosophical causality with Granger tests, but may be able to show that one series is a leading indicator for another. True causality has perplexed philosophers for millennia, so we will not settle the question here. Instead we will just present the most interesting results from many Granger tests for scientometric indicators.

Background

Scientometric models are similar to econometric ones. A nation's innovation establishment can be considered to be an economic system that needs inputs of resources like labor and capital to produce outputs such as products and exports. System inputs and outputs can be measured using indicators. Figure 1 shows the relations between the system model and these indicators. This is a simplified linear model of a more complex situation. In reality there are feedback loops--e.g., an overall one that shows that sales of products can provide resources for investments in R&D.

Previous cross-country analysis showed that there is a strong correlation between inputs and intermediate indicators like papers. Leydesdorff (1990) regressed world share of publications in the SCI as output on GERD as an input. Shelton (2006) identified national inputs most important in encouraging papers. His model suggested that changes in the GERD share have been the driver of national changes in paper share, which can account for the rise of China since 2001 (Jin & Rousseau, 2005; Shelton & Foland, 2010). Later, the models were refined

using components of GERD as explanatory variables (Leydesdorff & Wagner, 2009). Similar models showed that government investments in R&D and higher education spending on R&D (HERD) were especially effective, helping to explain Europe's passing the U.S. in papers during the 1990s (Foland & Shelton, 2010). Conversely, the industrial component of GERD was shown to be more effective in encouraging patents (Shelton & Leydesdorff, 2012). Here these methods are applied to high-technology (HT) outputs as an overall measure of the success of a national innovation enterprise. The preliminary cross-country analysis (Shelton & Fadel, 2014) raised questions about the direction of causality, so a longitudinal approach for time series for individual countries has now been added, using the Granger test.

Such analysis is becoming more common in scientometrics, but sometimes with limited results. After considerable effort, Vinkler (2008) found no significant link between economic performance and research. Peng (2010) found some causality between R&D expenditure and GDP in China, but it is not clear that his series had the required stationarity. LC Lee, Lin & YW Lee (2011) used Granger testing of whether research papers can be said to cause GDP output—aggregated by regions. One result was that there is mutual causality between research and economic growth in Asia, but the results are not so clear in the West. Inglesi, Chang & Gupta (2013) tried Granger testing between research papers and economic growth in Brazil, Russia, India, China, and South Africa (the "BRICS"), which mostly failed to demonstrate causality, except for some positive results for India. Inglesi, Balcilar & Gupta (2014) got more positive results for the connections between U.S. paper output and GDP.

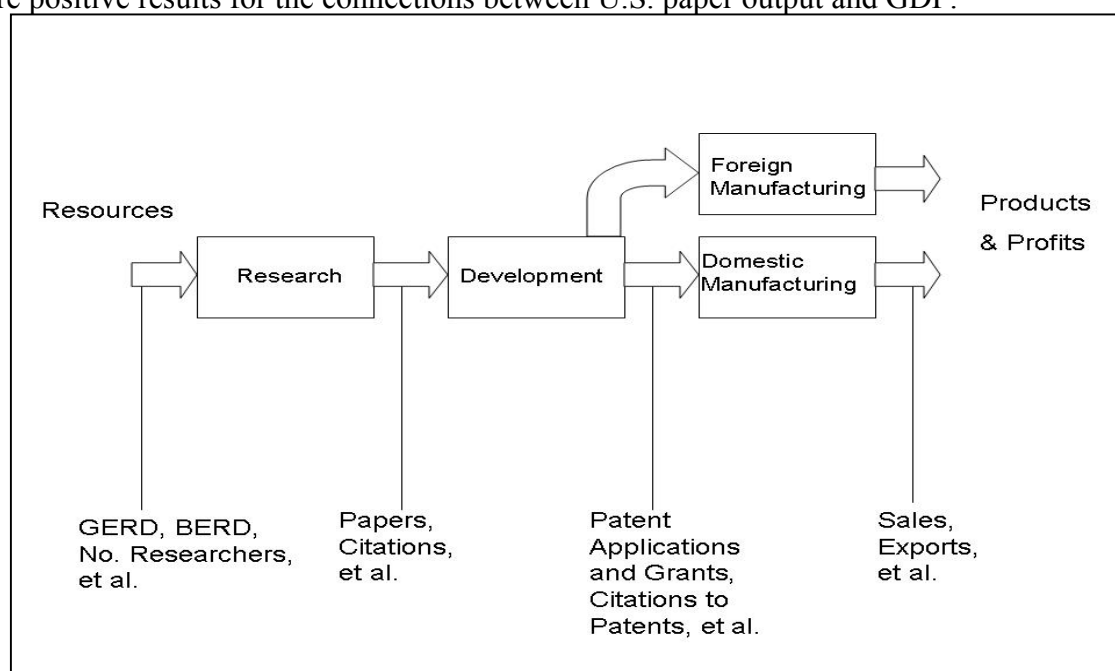


Figure 1. Linear model of an innovation enterprise with some indicators.

While there are some economic papers on factors that best explain *overall* international trade, there are relatively few that focus on the high-technology sector. One economic analysis of whether a country's high-tech exports (as a share of its overall exports) could be explained by R&D investment and country size was done by Braunerhjelm and Thulin (2008). They used the OECD data for 19 countries during 1981-1999. From their economic model, they concluded that overall R&D investment was significant.

Tebaldi (2011) used panel data to analyze factors that are most explanatory of high-technology trade. This approach adds data from more than one year to the usual cross-country analysis. Human capital, inflows of foreign direct investment, and openness to international trade were found to be the most significant of the factors he analyzed.

Data

Indicators like counts of papers and patents come from familiar sources like the SCI (Thomson Reuters 2015), (NSB 2014), and (OECD 2015). They provide insight into the success of national innovation enterprises. However, they are distant proxies for some of the quantities that the public cares most about: jobs, strength of their national economy, and survival of national industries. One scientometric measure of innovation that comes closer to these concerns is the performance of high-technology (HT) industries. Data on HT exports have been compiled on a cash basis for decades by the OECD (2015) in its Main Science and Technology Indicators series. However, this measure of industrial output does not capture the nuances of where manufacturing really takes place. For example, the Apple iPad is assembled in China, but most of its components come from Japan, the U.S. and elsewhere (Xing, 2012). Recently a new dataset has been jointly developed by the OECD and the World Trade Organization for manufacturing output on a value-added basis, which avoids double-counting of imported components. This more accurate data, as summarized in (NSB, 2014), allows development of much-improved models that tie these key outputs to inputs like R&D investment. Figure 2 shows some national time series for this measure of HT manufacturing output. Forecasts show that China will soon take the world lead as the U.S. and Japan move final assembly of HT products to China. (Similar graphs for HT exports on a cash basis showed China taking the world lead in 2005.) The Europeans, especially the Germans, seem to have done less of this "off-shoring." There have obviously been big changes in the last decade, and scientometric models might provide insight on why, and what governments might do to respond.

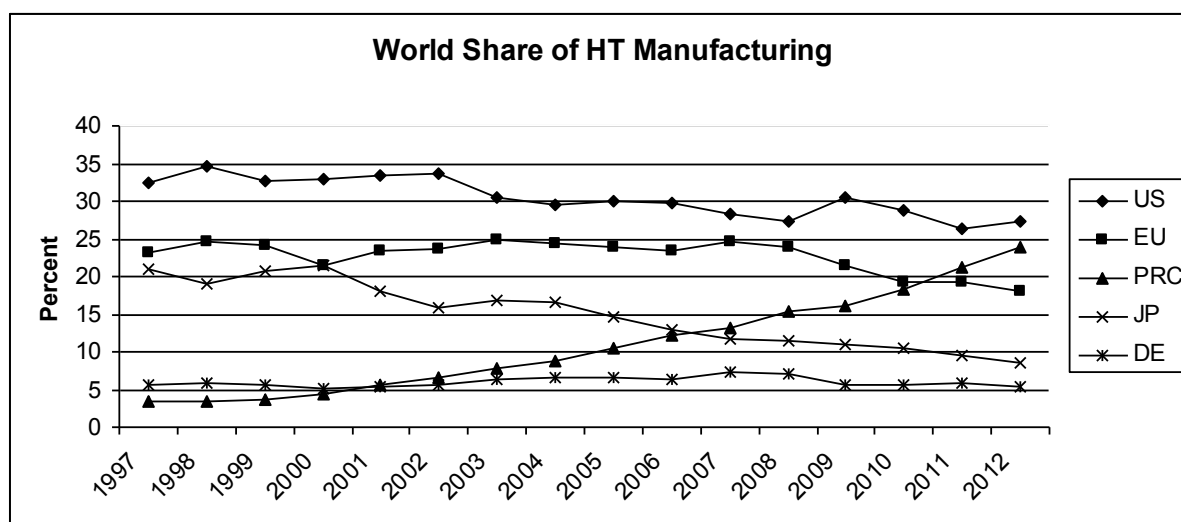


Figure 2. World share of manufacturing of high-technology products, on a value-added basis, for the United States, European Union (28), People's Republic of China, Japan, and Germany.

Causality Methods

Cross-country correlations over the countries in the OECD database are well known. Granger testing can be illustrated by revisiting the key results from Foland & Shelton (2010). That paper provided evidence that the EU passed the U.S. in papers in the mid-1990s because of a U.S. shift in research funding from government to industry, which was less effective in producing papers. At the time, this argument was based on cross-country correlations, and visual inspection of the U.S. and EU15 paper curves, which were very similar to their government GERD (GG) curves, just lagged by a couple of years. Granger testing can now add some quantitative evidence to this conclusion. First the series passed the ADF tests on the

data from 1988 to 2002, once second differences were calculated. The resulting Granger significance probabilities are in Table 1; bold entries are significant ($p < 0.1$).

Table 1. Significance probabilities for Granger tests of Government GERD component (USGGFF) causing papers (USPFF) on the NSI CD—or the reverse. FF means second difference. The “→” symbol means “Granger causes.”

<i>Model Order (k)</i>	<i>USGGFF → USPFF</i>	<i>USPFF → USGGFF</i>
1	p = 0.095	p = 0.52
2	p = 0.041	p = 0.19
3	p = 0.092	p = 0.73

Thus the government GERD indicator can be said to “Granger cause” papers in the U.S. in this time interval. The most significant result was for a model order of two years, and there was no significant reverse causality. This provides additional evidence that relative changes in the Government GERD component led to the EU becoming much more efficient than the U.S. in producing papers, and led to its passing the U.S. in the mid-1990s to become the world leader in this indicator.

The Granger test has low power, that is, it often does not find significant results, particularly when the sample size is small. The sample size for Table 1 is only $N = 15$, preventing the use of higher model orders, so it is fortunate that some definitive results were obtained. To seek more definitive results, longer series were extracted for US, EU15, Japan, Netherlands, and Turkey from the Web of Science and the OECD for 1980 – 2012 where possible. After the second differences necessary for stationarity, this resulted in $N = 30$ samples for 1982-2012.

One experiment investigated whether total GERD (using constant \$ and PPP weights) could be said to cause papers in the WoS (articles, letters, and reviews), with whole counts from the SCI-E and SSCI indexes. The results showed that U.S., EU15, and Japanese papers were indeed Granger caused by their national GERD with the significance probabilities in Table 2. None showed reverse causality. It did take a much higher model order to demonstrate Japanese causality. It was not possible to demonstrate significant results for the Netherlands or Turkey.

With these longer series, there is also the possibility that structural changes may take place over years. Sometimes a sliding window is used to examine shorter intervals within a longer one (Inglesi, Balcilar & Gupta, 2014). Here an auxiliary analysis simply examined the most recent years 2000 – 2012 ($N = 13$). The U.S. still exhibited Granger causality with the best result of $p = 0.012$ for a model order of $k = 2$. However, the other four country results for this shorter interval were not significant.

Table 2. Significance probabilities for Granger tests of GERD (G) causing papers (P) in the WoS (or the reverse) for 1983-2012. All used second differences.

<i>Order (k)</i>	<i>USG → USP</i>	<i>USP → USG</i>	<i>EUG → EUP</i>	<i>EUP → EUG</i>	<i>JPG → JPP</i>	<i>JPP → JPG</i>
1	0.0067	0.53	0.41	0.67	0.30	0.65
2	0.0024	0.89	0.53	0.94	0.53	0.53
3	0.013	0.90	0.76	0.82	0.54	0.54
4	0.034	0.91	0.085	0.92	0.54	0.79
5	0.10	0.86	0.14	0.96	0.064	0.80
6					0.0029	
7					0.0090	
8					0.011	

A similar test for Patent Cooperation Treaty (PCT) applications (OECD, 2014) in the U.S. was not so conclusive. Only for a model order of $k = 6$, could it be said that GERD Granger caused PCT patents, with $p = 0.09$. There was no reverse causality, however.

Another experiment tried to confirm a finding from Foland & Shelton (2010), that higher education spending on R&D (HERD) was closely associated with more papers. The dataset again included the U.S., EU15, the Netherlands, Japan, and Turkey, for the data range 1988-2002. Significant results were obtained only for the last two countries (Table 3). It was necessary to use fairly large model orders for Japan. The series passed the ADF tests with second differences, and there was no reverse causality for these model orders. Thus it can be said that, in Japan and Turkey at least, HERD Granger caused papers in these years. This might be useful for professors in those countries to mention in their battles for more funding.

Table 3. Does higher education spending Granger cause scientific papers?

<i>Model Order (k)</i>	<i>Japan HERD $\rightarrow$ Japan Papers</i>	<i>Turkey HERD $\rightarrow$ Turkey Papers</i>
1	$p = 0.56$	$p = 0.24$
2	$p = 0.82$	$p = 0.049$
3	$p = 0.37$	$p = 0.12$
4	$p = 0.090$	$p = 0.21$
5	$p = 0.016$	$p = 0.30$

Correlations for the Value-Added HT Manufacturing Indicator

Simple correlation over the 40 or so countries in the database of input resources in (OECD, 2014) can provide insight into which investments might be most productive in encouraging HT exports and manufacture. However, since many indicators simply increase with the size of the country, it is necessary to find explanatory variables whose correlations are much greater than those for measures like population or GDP. Furthermore, the U.S. and China are outliers; it is necessary to either omit them, or use log measures, if the contributions of smaller countries are to affect the results.

Table 4 from Shelton & Fadel (2014) shows the coefficients of determination (R^2) for two measures of performance of national HT industries with a number of explanatory or independent variables. For both measures, business expenditure on R&D (BERD) is best, with gross expenditure on R&D (GERD) not far behind. The correlations are far better for the new value-added data for HT manufacturing in the last column, than for the earlier exports on a cash basis. Indeed a quite accurate regression model can be constructed for this case (Equation 1), where NM9 is HT manufactures and BN9 is BERD, both in current dollars in 2009. Figure 3 shows the scattergram for this model.

$$\log \text{NM9} = 0.385 + 0.944 \log \text{BN9} \quad (R^2 = 84.1\%) \quad (1)$$

One would expect that there would be a delay between R&D investments and downstream benefits. For some indicators like patent grants, models that incorporate these delays can be more accurate (Shelton & Monbo, 2012). Here, correlations do not change much with lags, thus they did not improve the models enough to warrant the increased complexity. To see if a multiple linear regression would improve the model, a step-wise regression on HT manufacturing in 2009 was performed using the nine independent variables in Table 4. None of the other variables was significant in a multiple regression, once BERD was included as an explanatory variable, making a simple univariate regression without lags reasonable.

Table 4. Coefficients of determination (R^2 in %) of HT exports and overall HT manufacturing with explanatory variables in 2009. Uses log scales. More recent data downloads produce somewhat different correlations, and the values are sensitive to missing data points.

	<i>Exports (Cash Basis)</i>	<i>Overall Output (Value-Added)</i>
Papers SCI	41.7	71.0
Patents Triadic	48.8	69.9
Patent PCT Apps	34.3	61.5
GERD	44.8	79.8
BERD	49.0	84.5
Researchers	26.2	61.4
Business Researchers	29.3	71.6
Size GDP	27.3	56.9
Size Population	13.1	34.3

Despite the precision of the regression model in Equation (1), however, there is an alternate explanation for the trends of HT manufactures in the last decade. Could it be that HT manufacturing causes R&D investment, instead of the reverse? Indeed, it is the income from these sales that does provide some of those resources. OECD states that it picked the sectors for inclusion in the HT set precisely because these industries invest an extraordinary fraction of their income in R&D. And these correlations are too good to be true for BERD solely causing HT manufacturing--there are simply too many other factors that must also contribute. There have been frequent news accounts of Western and Japanese firms moving manufacturing to China and other low wage countries to increase their profits. China was also favored because its vast market offered potential for huge growth in HT sales.

This alternate explanation brings into question the efficacy of a nation increasing its HT manufacturing by encouraging greater business investment in R&D. It is possible that the results might be disappointing if the executives of the HT companies still prefer to locate the manufacturing abroad, the top path in Figure 1, so that some other nation reaps the benefits of the sales of HT goods. A policy remedy that addresses both explanations would be more likely to succeed. R&D investment policies could be coupled with trade policies that encourage location of manufacturing where the investments were made.

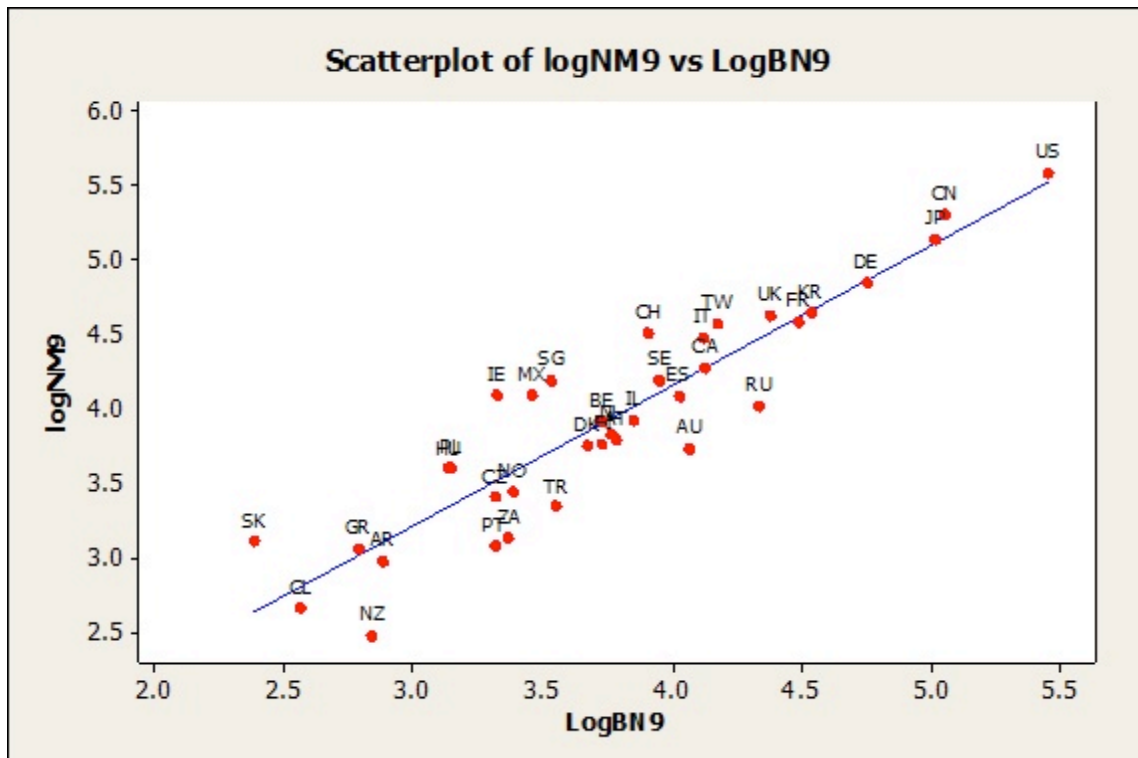


Figure 3. Scattergram of overall high-technology manufacturing vs. business expenditure on R&D in 2009. The cluster in the center contains BE, DK, IL, FI, and NL. HU and PL also overlap.

Further both manufacturing and BERD could be the results of an exogenous variable, some underlying third series. For example many of them seem to be closely tied to recent perturbations of the business cycle over the 1998 - 2011 data range available.

Causality Results for Value-Added HT Manufacturing

Table 1 shows that BERD has the highest correlation with HT manufacturing, so it will be analyzed first. Overall results for the sum of all countries in the OECD database were not significant. Findings for those individual countries with significant results are in Table 5. All are for model order $k = 1$, but orders up to $k = 3$ do not add countries to the list. Both series use current dollar values, and BERD used PPP weighting. The data ranges from 1999-2012.

Table 5. Does BERD Granger cause HT manufacturing (Mfg), or the reverse? Entries are significance probabilities; $p < 0.1$ is significant (bold type).

<i>Country</i>	<i>Mfg → BERD</i>	<i>BERD → Mfg</i>
Korea	0.21	0.097
Hungary	0.16	0.0013
Romania	0.57	0.023
PRC	0.025	0.32
Canada	0.019	0.43
Germany	0.016	0.19
Russia	0.060	0.54
Finland	0.0014	0.010

Of the some 24 countries with complete OECD data, 15 passed both ADF tests using second differences. The entries in bold type are the only ones that were significant from the Granger tests. While these results do not settle the question, they do show that (Granger) causality can indeed run in either direction for these indicators. Policymakers in Korea, Hungary, and Romania could benefit from knowing their country's business R&D investment did Granger cause its HT manufacturing output in these years, and may want to encourage more of this virtuous cycle. (Taiwan also showed this direction of causality for its available data from 2000-2012, using model order $k = 2$.) Chinese, Canadian, German, and Russian policymakers might be pleased to find that their country's HT manufacturing output Granger caused more BERD investment. Those in Finland would probably not find bi-directional causality very useful.

The second highest correlation in Table 1 was with overall GERD. As expected, these results were not as conclusive as those for the BERD component. Of the some 40 countries in the OECD Group, 30 had complete data. Of these 13 passed the ADF test for stationarity for both time series, using second differences. Using $k = 1$, only Hungary and Korea showed positive results ($p = 0.0029$ and $p = 0.069$, respectively). In the reverse direction of Mfg causing GERD, only Canada and Germany showed significant results ($p = 0.026$ and $p = 0.0075$ respectively). The Slovak Republic showed bi-directional causality with $p = 0.091$ for GERD causing Mfg and $p = 0.025$ in the reverse direction. These results seem to show that the higher correlation of BERD with manufacturing is necessary to get more definitive results.

BERD and GERD are not always thought of as scientometric indicators, though. What can be said about causality of HT manufacturing for traditional intermediate scientometrics indicators like papers and patents? Only a couple of countries had significant results for papers, but the PCT patent applications were more interesting (Table 6). Using second differences, the ADF tests showed that 29 countries of the 37 countries with data had both series stationary, and 10 countries, plus the EU as a whole, showed Granger causality. The results are for order $k = 1$, except for Denmark and the Czech Republic where $k = 2$. Two countries had bidirectional causality: Germany and the Netherlands.

Table 6. Do PCT international patent applications Granger cause HT manufacturing, or the reverse? Entries are significance probabilities; $p < 0.1$ is significant (bold type).

<i>Country</i>	<i>Patents→Mfg</i>	<i>Mfg→Patents</i>
EU28	0.060	0.14
Austria	0.036	0.24
Belgium	0.046	0.24
Canada	0.060	0.15
Czech Republic	0.055 (k = 2)	0.54
Denmark	0.012 (k = 2)	0.78
Korea	0.063	0.92
New Zealand	0.0064	0.11
Switzerland	0.014	0.40
Germany	0.0014	0.0047
Netherlands	0.050	0.055

So, there are quite a few countries where it can be said that their patenting activity Granger causes HT manufacturing output. This connection was suggested by the correlation results in Table 1, of course. There are good physical reasons that make this causality plausible, but the results do *not* imply that a national initiative to file more PCT applications would necessarily

result in more manufacturing. The Granger tests do add quantitative evidence that investments in science and technology indeed bear fruit in outputs that the public cares about.

Conclusions

For further work, statistical testing for causality can enrich study of the connections between scientometric indicators, and there are many others. However, the Granger test often fails, even when strong cross-country correlations exist and there are good physical reasons to suspect causality. There are other tests, like Toda & Yamamoto (1995), which can be employed. And more sophisticated data analysis might also help: other methods of de-trending, sliding windows for long series, panel data, et al. As always, one needs to be cautious of spurious results from data mining; running many tests is likely to turn up some positive results by chance.

The results here show that GERD did Granger cause papers and patents for the U.S., which is probably true for some others as well. This quantitative evidence bolsters the case that R&D funding is important for the success of a nation's science. In particular, the U.S. has a goal of maintaining its science leadership, but is rapidly falling behind in the funding race with China. In a rare good year, the U.S. increases its GERD by a real 3%; Chinese GERD has been increasing by more than 15% annually for decades.

New data on value-added manufacturing outputs provides quantitative insight on which inputs can be most effective in encouraging high-technology industries. Not surprisingly, there is a strong connection between such success and investments in R&D, particularly by the business sector. In countries where this can be demonstrated to be a cause of these successes, governments might wish to adopt policies, such as tax incentives, which can encourage such investment. Intermediate indicators like patents can also be good explanatory variables, showing quantitatively that traditional scientometric measures indeed provide useful information about outputs that directly affect a nation's prosperity.

Of course there are many other benefits of science and technology beyond the manufacture and sale of the HT products considered here. Science can lead to better healthcare, cleaner air and water, solutions of problems like global warming, improved communications that allow more extensive cooperation and collaboration, and many others. Most of these benefits can accrue to everyone, regardless of their nationality. Even in the competitive analysis of national market share of HT manufactures considered here, one should not lose sight of the overall performance of the sector. Worldwide sales have almost doubled over the last decade with only a slight pause during the Great Recession, reaching over \$1.5 trillion in 2012. This growth has created millions of new jobs and a cornucopia of wonderful new products most people can enjoy--the ubiquitous cell phone has provided the first rapid communications in some of the poorest countries.

Acknowledgments

This work was partly funded by NSF cooperative agreement ENG-0844639. These findings do not necessarily reflect the views of NSF.

References

- Braunerhjelm, P. & Thulin, P. (2008). Can countries create comparative advantages? R&D expenditures, high-tech exports, and country size in 19 OECD countries 1981-1999. *International Economic Journal*, 22, 95-111.
- Foland, P. & Shelton R.D. (2010). Why is Europe so efficient at producing scientific papers, and does this explain the European Paradox? *11th International Conference on S&T Indicators*, Leiden.
- Granger, C.W.J., (1969). Investigating causal relations by econometric models and cross spectral methods. *Econometrica*, 37(3), 424-438.

- Inglesi-Lotz, R., Balcilar, M., & Gupta, R. (2014). Time-varying causality between research output and economic growth in US. *Scientometrics*, 100, 203-216.
- Inglesi-Lotz, R., Chang, T. & Gupta, R. (2013). Causality between research output and economic growth in BRICS. *Quality & Quantity*.
- Jin, B. & Rousseau, R. (2005). China's quantitative expansion phase: Exponential growth, but low impact. *Proceedings of the 10th International Conference on Scientometrics and Informetrics*, Stockholm.
- Krugman, P. (1991). *Geography and Trade*, Cambridge: MIT Press.
- Lee, L.C., Lin, P.H., Chuang, Y.W., & Lee, Y.Y. (2011). Research output and economic output: a Granger causality test. *Scientometrics*, 89, 465-478.
- Leydesdorff, L. (1990). The prediction of science indicators using information theory. *Scientometrics*, 19, 297-324.
- Leydesdorff, L., & Wagner, CS. (2009). Macro-level indicators of the relations between research funding and research output. *Journal of Informetrics*, 3(4), 353-362.
- NSB. (2014). *Science and Engineering Indicators 2014*. Arlington: National Science Board.
- OECD. (2015). *Main Science and Technology Indicators*. Retrieved January 1, 2015 from: http://stats.oecd.org/Index.aspx?DataSetCode=MSTI_PUB
- OECD. (2014). *Trade in Value-Added*. Retrieved February 23, 2014 from: <http://www.oecd.org/sti/ind/49894138.pdf>
- Peng, L. (2010) Study on the relationship between R&D expenditures and economic growth of China. *Proceedings of the 7th International Conference on Innovation and Management*. 1725-1728.
- R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved on Dec. 1, 2014 from <http://www.R-project.org/>
- Seth, A (2007) Granger causality. Retrieved April 17, 2015 from http://www.scholarpedia.org/article/Granger_causality
- Shelton, R.D. & Fadel, T.R. (2014). Which scientometric indicators best explain national performance of high-tech outputs? *15th Collnet Conference*, Ilmenau, Germany.
- Shelton, R.D. & Monbo, S.D. (2012). Input-output modelling and simulation of scientometric indicators: A focus on patents. *Proceedings of the 17th International Conference on S&T Indicators*, pp. 756-767. Montreal.
- Shelton, R.D. & Leydesdorff, L. (2012). Publish or Patent: Bibliometric Evidence for Empirical Trade-offs in National Funding Strategies. *Journal of the American Society for Information Science and Technology*, 63(3), 498-511.
- Shelton, R.D. (2006). Relations between national research investment and publication output: Application to an American paradox. *Scientometrics*, 74(2), 191-205.
- Shelton, R.D. & Foland, P. (2010). The race for world leadership of science and technology: Status and forecasts. *Proceedings of the 12th International Conference on Scientometrics and Informetrics*, pp. 369-380.
- Thomson Reuters (2015) Web of Knowledge. Retrieved on January 1, 2015 from <http://wokinfo.com/>
- Tebaldi, E. (2011). The determinants of high-technology exports: A panel data analysis. *Atlantic Economic Journal*, 39:343-353.
- Toda, H.Y. & Yamamoto, T. (1995) Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 66, 225-250.
- Vinkler, P. (2008). Correlation between the structure of scientific research, scientometrics, and GDP in EU and non-EU countries. *Scientometrics*, 74, 237-254.
- Xing, Y. (2012). *The PRC's High-Tech Exports: Myth and Reality*, ADBInstitute Working Paper #357. Retrieved January 13, 2014 from: <http://www.adbi.org/working-paper/2012/04/25/5055.prc.high.tech.exports.myth.reality/>

Scientific Production in Brazilian Research Institutes: Do Institutional Context, Background Characteristics and Academic Tasks Contribute to Gender Differences?

Gilda Olinto¹ & Jacqueline Leta²

¹*gilda@ibict.br*

Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Rua Lauro Muller, 455 - 4º andar, CEP 22290 – 160, Rio de Janeiro (–Brazil)

²*jleta@bioqmed.ufrj.br*

Universidade Federal do Rio de Janeiro (UFRJ), Av. Brigadeiro Trompowsky s/ nº, Prédio do CCS, Bloco B – sala 39, CEP 21941-590, Rio de Janeiro (Brazil)

Abstract

Despite the recent changes that occurred in the Brazilian science, this field is still strongly anchored on male figures, as it happened at the beginning of its institutionalization. This paper detaches the contribution of Brazilian Research Institutes for the development of Brazilian science and the importance of contextual, background and academic tasks involvement in scientific production in those institutes, giving special attention to gender differences. Data from government graduate programs evaluation forms were obtained for the analyses presented here which take into account all professor-researchers - 890 women and 1,470 men - affiliated to 72 graduate programs under the responsibility of 31 Brazilian Research Institutes (BRI), the majority of which supported by the Federal Government. The main findings include: women are a minority in those institutes, are concentrated in the health and biological sciences, show higher scientific production than their male colleagues, especially in journal articles and among those involved in highly evaluated graduate programs. We believe the set of results presented in this paper may contribute to a better understanding of women's participation not only in BRI, which are dedicated to specific scientific areas, but also in Brazilian science in general and so contribute to gender governmental policy.

Conference Topic

Country level studies

Introduction

The process of science institutionalization in Brazil started about a century ago, when in Europe and in the USA this activity was already structured, both in science academies and in research institutions. One of the first steps contributing to this process in Brazil was the creation, in 1900, of the Federal Serotherapy Institute at Manguinhos, in Rio de Janeiro (which was afterwards named Instituto Oswaldo Cruz), considered the first Brazilian Research Institute to win international recognition (Weltman, 2002). In the following decades, the first public universities were created, as the University of Brazil (later renamed Universidade Federal do Rio de Janeiro), founded in 1920, and the University of São Paulo, in 1934. However, only in the nineteen fifties, with the creation of the first agencies for the promotion of scientific development in the country, this process advanced significantly: CAPES assumed the responsibility of structuring and monitoring graduate programs (Masters and Doctorate), throughout the country, while the other agency, the CNPq assumed the task of promoting scholarships and research projects.

Considering the above mentioned initiatives, it is possible to say that, in the second half of the twentieth century, one witnesses a strong governmental effort towards structuring scientific institutions, and also an induced and spontaneous expansion of graduate programs. In 2010, three decades later, the country already counted with an extensive system of S&T, including: 83,170 doctors-researchers, 64,588 students enrolled in doctorate courses, 2,840 graduate programs, 27,523 research groups, and 452 research institutes and universities throughout the

country (MCTI, 2014). The effort to train and qualify S&T human resources, build up and modernize the infrastructure of research institutions and, more recently, create legal tools to allow the increase and maintenance of science funding, resulted in an outstanding growth of scientific output in the years 2000, especially output in journals indexed by international bibliographic databases (Regalado, 2010; Leta et al., 2013).

It is important to point out that such growth is also result of a combination of factors, besides the previously mentioned ones. Among these factors, the following could be mentioned: (1) the inclusion of Brazilian journals in databases, which resulted in an expressive growth of Brazilian production in international bases in the last few years (Leta, 2012); and (2) the creation of evaluation mechanisms of graduate programs, which stimulate and reward output in journals, mainly in international journals (Mugnaini & Sales, 2011). About this last aspect, it is important to highlight that graduate programs - which cover all areas of knowledge and a great part of the institutions of higher education and research, especially those of the public sector - became the leading stronghold of Brazilian science. Thus, policies and evaluation mechanisms directed to these programs are reflected in Brazilian scientific outputs and outcomes.

The institutionalization, growth and international recognition of Brazilian science have not promoted significant changes in aspects of scientific stratification, more specifically an equalitarian representation of men and women in scientific activities. Although the last decades have witnessed a significant growth in the number of women in the country's academic and scientific fields – in higher education, in graduate programs and as professors and/or researchers at universities and research institutions (INEP, 2007) – they are still a minority in several areas, in higher academic levels and in administrative functions of higher prestige (Olinto, 2011; Gauche, Verdinelli & Silveira, 2013). This scenario, although not exclusive of Brazilian scientific field, calls attention to the fact that, in face of the many recent changes that occurred in the country's science, this field is still strongly anchored on male figures.

Many factors support the maintenance of this scenario in Brazil and in the world, where women are excluded of certain areas, a phenomenon known as horizontal gender segregation, and they do not advance in their careers, a phenomenon known as the vertical gender segregation (Shienbinger, 2001). In a previous study (Leta et al., 2013), considering the symbolic value of different academic tasks that are part of the academic career, the hypothesis posed was that female Brazilian scientists would be involved in tasks of lesser prestige and, consequently, would be less productive and advance less in their careers than their male peers. We inquired into this issue examining productivity and involvement in academic tasks of the population of over 52,000 professor-researchers who participated in Brazilian graduate programs (our unit of analysis was each professor-researcher linked to a Brazilian graduate program, and whose academic characteristics and performance are yearly included in evaluation forms provided by the federal government). This study revealed a higher participation of men in articles published in annals of events, but major differences between male and female professors-researchers were not observed. Even though it may be considered positive the fact that both sexes have an equal share of academic-scientific tasks, the population analyzed in the mentioned study was very heterogeneous. Subtle differences were found, however, when the analysis considered the area of graduate work in which the professor-researcher was linked to. The health area was the closest one to our hypothesis: women tend to get more involved in activities of lesser prestige, like teaching graduate courses, and less involved in activities of higher prestige, like publishing in journals. Academic area and the nature of the institution are some aspects, among others, that may have an impact in the characteristics and the amount of scientific output of both men and women. In order to reduce diversity, in the present study, the focus turned to the participants of

graduate programs who are affiliated to Brazilian Research Institutes. The central question of this study is: how do gender differences in scientific performance are related to the characteristics of the academic and institutional context, as well as the involvement in several academic tasks of professor-researchers in graduate programs of Brazilian Research Institutes?

Research Institutes and Women

The largest part of the Brazilian Research Institutes belongs to the public sector and is linked to the Ministry of Science, Technology and Innovation (MCTI). Among the oldest is the National Observatory, founded in 1827, in the city of Rio de Janeiro. Presently there are thirteen other Research Institutes linked to the MCTI, the majority directed towards research in exact sciences and engineering. Other ministries also maintain Research Institutes, as the Ministry of Agriculture, responsible for Embrapa, created in 1973 with the purpose of developing research in agriculture; the Ministry of Health is responsible for the Brazilian National Cancer Institute (INCA), founded in 1961, and for the Oswaldo Cruz Foundation (at present – Fiocruz), created in 1900.

Until recently, women's presence and contribution at Research Institutes was poorly explored as a research topic in studies about gender and science. Among a few recent studies, the one by Brito Ribeiro (2011) inquired into the distribution of male and female researchers at Research Institutes linked to the MCTI in two career functions: researcher and technologist. This author points out to the small proportion of women in those institutes: about 30% in both types of careers. Nevertheless, that fraction still decreases substantially when the research areas of these institutions are considered. In the Brazilian Center of Research in Physics, for instance, there are only 17% of women in those two careers. The author also presents data about the distribution of men and women in higher prestige posts at these institutions, like presidency and boards of directors: out of 362 senior administrators, only 36 (10%) were occupied by women in 2010, a clear indication of vertical gender segregation. A more thorough analysis was done recently taking into account 571 researchers, with doctor degrees, affiliated to Fiocruz (Rodrigues, 2014), an institution that plays a central role in health research in the country. This author points out that male researchers have a *per capita* output quite superior to that of female ones. A different situation is found in Fiocruz, however, when the analysis focuses on administrative positions. Differently from other Research Institutes, especially those oriented towards exact sciences and engineering, Fiocruz is concerned with gender equity, and thus started a Pro-Equity Gender Program in 2009. This initiative might explain the large number of women in administrative positions in this institution. In 2013, out of 768 administrators with salary bonus, 382 (49.7%) were women, which is close to parity. However, women are still an absolute minority occupying the highest prestige posts, as president and directors.

The scenario previously described is shared by Research Institutes of other countries. One of the most prominent Research Institutes in the world, the Massachusetts Institute of Technology, has recently published a study on gender equity in the institution. Compared with previous studies (1999 and 2002), it showed major advances in two Schools. In the School of Science and School of Engineering, particularly, "the number of women in faculty increased significantly (from 30 to 52 in science and 32 to 60 in engineering) and in both schools women now hold several senior administrative positions" (Gillooly, 2011). However, despite these advances, women are still a minority, especially among those that occupy positions of higher prestige and salary, as tenured faculty members, of which women represent only 15% and 12% in the two schools, respectively. At the Centre National de la Recherche Scientifique (CNRS), the largest Research Institute in France, a country with a solid tradition in science and a pioneer in actions and policies that benefit women, Hermann

& Cyrot-Lackmann (2002) observed that women represent from 22% to 38% of the total CNRS's researchers and, what seems to be more significant, 31% of the research directors are in the highest prestige positions. Yet, as seen in the MCTI Institutes in Brazil, at the CNRS in France, this representation also varies according to the area of study: in Physical & Mathematical Sciences and Engineering Sciences only 12% and 9%, respectively, are women; and in Life Sciences, 28% of the research directors are women.

Different theories and models are considered by the literature to explain the phenomenon of female segregation in science and they include personal, biological, cultural, social and institutional aspects; and empirical studies based on these theories and models usually point out to gender imbalances favoring men (Barrios, 2013; Epstein, 2007; European Commission, 2009; Fox, 2005; Long, 1992; Meulders et al., 2010; Prpic, 2002).

The present focus on gender differences in institutional contexts suggests that male researchers would show better performance in different academic tasks and also present greater scientific production, like publishing in prestigious journals. Rewards for better performance would include the occupation of prestigious posts. Such arguments allow one to bring about the concept of scientific capital, proposed by Bourdieu (2003): a kind of symbolic or tacit capital, which opens opportunities and promotes recognition and which would tend to help perpetuate gender differences in science. Researchers with higher rates in publications and with high involvement in prestigious academic-scientific tasks accumulate scientific capital and, in a "snow ball" feedback effect, would tend to keep to themselves positions of higher academic prominence. In an opposite movement, researchers with less involvement in the more valued activities accumulate less scientific capital and would tend to be less involved in the more valued tasks, as well as to have a greater burden of less valued tasks, as, for instance, teaching assignments. Considering this model, the present study intends to investigate the relation between gender, academic background, institutional context, including the involvement in academic tasks, and scientific output of professor-researchers affiliated to the BRI.

Data collection and method

This study uses the documental analysis technique applied to information retrieved from three pre-established PDF forms with information used in the 2009 national evaluation of graduate programs (CAPES, 2013). Information provided includes aspects of academic and scientific performance as well as personal and academic characteristics of 52,294 professor-researchers affiliated to 2,247 graduate programs. Since a key characteristic, the professor-researcher's gender, was not included in CAPES' forms, a series of strategies was developed to allow for this classification (Leta et al., 2013).

For the present study, we have selected a subset of the 2009 original population and took into account information about all professor-researchers affiliated to 72 graduate programs under the auspices of 31 Brazilian Research Institutes (BRI), which were classified by us in three main groups: (1) supported by funds from the Federal government (Public/Federal), (2) supported by funds from State governments (Public/States) and (3) supported by the private sector (Private).¹

¹ First group: Brazilian Center of Research in Physics (CBPF), Centre of Nuclear Technology Development (CNEN/CDTN), Institute of Nuclear Engineering (CNEN/IEN), Institute of Radio Protection and Dosimetry (CNEN/IRD), Oswaldo Cruz Foundation (Fiocruz), Research Centre (FIOCRUZ/ CPqGM), René Rachou Research Centre (FIOCRUZ/CPqRR), Institute of Military Engineering (IME), Institute of Pure and Applied Mathematics (IMPA), Brazilian National Cancer Institute (INCA), National Institute of Metrology, Quality and Technology (INMETRO), National Institute of Research in the Amazon (INPA), National Institute for Space Research (INPE), National Institute of Industrial Property (INPI), Technological Institute of Aeronautics (ITA), Botanical Garden Foundation of Rio de Janeiro (JBRJ), National Laboratory for Scientific Computing (LNCC)

It is important to mention that not all BRI are included in this study since a few of them do not have a graduate program under their responsibility. Examples are Embrapa and IBICT, major research institutes in the areas of agricultural sciences/biology and information science, respectively. These Institutes do have graduate programs but they are organized in collaboration with public universities.

Once the BRI were identified and data cleaned, all information was exported to a matrix of SPSS (Statistical Package for the Social Sciences), version 12. The population of the study represented in this matrix, and focus of the analyses presented here, can be so defined: BRI professor-researchers who participated in graduate programs in Brazil in 2009 (N=2,362). Among the variables that characterize each professor-researcher are: (a) personal and academic characteristics of the professor-researcher (gender, S&T area and year of doctoral title), (b) characteristics of institution of affiliation/ graduate programs (economic sector, area and evaluation grade); (c) academic roles performed by each professor-researcher (graduate courses, graduate advising, banking participation, project leadership) and (d) publication output (journal articles, articles in Annals and other types of publications). For the classification of S&T area of the graduate programs, we utilized the categories considered by CNPq (2013).

Results

The analyses are presented in two main sections: (a) characteristics of the institutional context in which professor-researchers participate and aspects of his academic background and (b) academic tasks and the scientific output of the professor-researchers, with emphasis given to gender differences.

Characteristics of the Institutions and of professor-researchers background

Table 1 shows the distribution of the 2,362 professor-researchers according to three macro-characteristics of the graduate programs of the BRI to which these professionals are linked: the economic sector, the area of knowledge and the performance grade.

Considering the economic sector, data show that the greatest part of professor-researchers are linked to the institutions maintained by the Federal Government and very few of these professionals are active in programs belonging to private institutions: only 3%. These results are different from those obtained for Brazilian graduate programs considered as a whole, which showed that 55% of the institutions belonged to the federal government, 30% states government and 15% to the private sector (CAPES, 2014).

The distribution of professor-researchers according to the academic areas of the BRI graduate programs (which represent the areas of expertise of these professionals) is, however, more homogeneous, although it is clear that a massive number of professors are concentrated in two major groups: Engineering and Exact Sciences, in one hand, and in Health and Biological Sciences, in the other hand. These areas together absorb 80.3% of the professor-researchers in the BRI.

and National Observatory (ON). The second group: Nuclear and Energy Research Institute (CNEN/IPEN), Institute of Medical Assistance to the State Civil Servants (IAMSPE), São Paulo Institute of Biology (IBSP), São Paulo Institute of Botanic (IBT), São Paulo Institute of Fishery (IP), Institute of Ecological Research (IPÊ), São Paulo Institute of Technological Research (IPT), Pernambuco Institute of Technology (ITEP) and Institute of Zoology (IZ / APTA). Third group: Recife Centre of Studies and Advanced Systems (CESAR), Brasília Institute of Public Law (IDP), Latin American Institute of Research and Education in Odontology (ILAPEO) and Institute of Technology for the Development (LACTEC).

Table 1. Number and % of professor-researchers according to the economic sector, areas and grades of Graduate Programs from Brazilian Research Institutes – 2009.

<i>ECONOMIC SECTOR</i>	<i>N</i>	<i>%</i>
Public / Federal	1,933	81.8
Public / States	357	15.1
Private	72	3.0
Total	2,362	100
<i>AREAS</i>		
Engineering	489	20.7
Exact Sciences	476	20.2
Health Sciences	601	25.4
Biological Sciences	331	14.0
Human Sciences	71	3.0
Social Applied Sciences	14	0.6
Agrarian	31	1.3
Other/interdisciplinary	349	14.8
Total	2,362	100
<i>CAPES EVALUATION</i>		
Grade 2	38	1.6
Grade 3	356	15.1
Grade 4	623	26.4
Grade 5	693	29.3
Grade 6	489	20.7
Grade 7	163	6.9
Total	2,362	100

Table 2. Distribution (%) of professor-researchers from Brazilian Research Institutes according to academic areas and other characteristics by gender – 2009.

Contextual aspect	Percentage¹	
	Women	Men
Professor-researchers ²	37.7 (n= 890)	62.3 (n=1,470)
<i>ACADEMIC AREAS</i>	<i>%</i>	<i>%</i>
Engineering	8.5	28.1
Exact Sciences	10.8	25.9
Health Sciences	38.1	17.8
Biological Sciences	20.9	9.9
Other areas/interdisciplinary	21.7	18.4
TOTAL	100	100 ³
<i>OTHER CHARACTERISTICS</i>	<i>% yes</i>	<i>% yes</i>
Public / Federal	83.7	80.8
PHD before 2000	58.1	66.1
PHD abroad	16.4	30.0
Program with grade 2 to 3	14.5	17.9
Program with grade 5 – 7	59.0	55.8
Program with grade 6 to 7	20.6	31.9

Percentages calculated within each gender category. ² We were not able to attribute the sex of two professor-researchers. ³ Partial and total percentages provided by SPSS.

The final contextual aspect, presented in table 1, refers to the performance grade of the graduate programs issued by CAPES. These grades are recorded in a scale from 2 to 7, and the meaning of these assessments is: from grade 5 the program is considered to be at a good

level, able to participate in institutional programs etc. Grades 6 and 7 are assigned to programs of high performance, and some aspects that contribute to the assignment of these grades, besides scientific productivity, are institutional agreements as well as institutional exchange of researchers, professors and students. In table 1, it is also possible to observe that the great majority of professor-researchers participate in programs that received grades from 5 to 7.

The following Table 2 aims to identify gender differences in institutional affiliation and aspects of personal background of the professors/researchers in BRI.

It is possible to note that women represent less than 40% of this population (N=890), a fraction similar to the one obtained in a previous study which focused on professor-researchers of all graduate programs in the country (Leta et al., 2013). Data also show that women are predominant in the areas of Biological and Health Sciences, whereas men form a great majority in Engineering and Exact Sciences, which points to the phenomenon of horizontal segregation of gender, a characteristic also observed in Brazilian graduate programs in general (Leta et al., 2013).

Table 2 also presents other relevant information related to gender, calling attention to gender differences favoring men: a higher proportion of men show longer careers than women (which in fact might reflect the recent increase in women's entrance in scientific careers), relatively earn more degrees abroad and participate more in graduate programs of higher prestige.

Gender and scientific production of professor-researchers of Brazilian Research Institutes

Table 3 shows the distribution of men and women according to the number and the kind of published work in 2009 - articles in journals, complete works in annals of events and abstracts in annals of events.

Table 3. Distribution (%) of professor -researchers from Brazilian Research Institutes by sex and number of journal articles, annals full article and annals abstract – 2009.

Publication	Journal Article		Annals full Article		Annals Abstract	
	Women	Men	Women	Men	Women	Men
0	30.6	38.7	76.7	66.7	68.9	80.3
1-2	33.9	31.7	14.7	15.6	15.7	10.9
3+	35.5	29.6	8.5	17.7	15.4	8.8
Total	890	1,470	890	1,470	890	1,470

These results call attention to the high percentage of both men and women without any work published in 2009, particularly those with zero annals full article and annals abstract. This table also stresses the higher women's performance as far as journal articles are considered: a lower proportion of women are included among those with zero contribution to this kind of publication and a higher proportion of this gender group are among those contributing with one or two journal articles, and especially among those considered more productive: three or more articles. It is important to keep in mind that this is the kind of published work that has more value in the scientific field in general, and is also the kind of publication that contributes the most to the grades attributed to the graduate programs by Brazilian Agencies. In Annals, a type of publication that is highly valued in technological fields, as Engineering, it is possible to see an alternate pattern between men and women: men with better performance in annals full articles and women in annals abstracts.

Scientific production is influenced by a large number of factors, including the academic area, years of academic experience (Bonaccorsi & Daraio, 2003), education abroad (Velema, 2012), etc. Table 4 presents the publication mean of the different types of publications of the BRI professor-researchers by gender, as well as by gender controlled by the above-mentioned factors – area, experience and education abroad –, and also the CAPES grade of the program, a particular aspect in the Brazilian scientific area.

Taking into account the general mean performance and gender, table 4 also shows, as in table 3, that women outperformed men in BRI in 2009 in mean number of journal articles (women published a 2.51 and men 2.12 articles, mean results with similar standard deviation) and the mean number of annals abstract (W=1.14 and M=0.75), while men attained higher means of annals full articles (W=0.74 and M=1.48). With these results, and considering the higher academic value attributed to publication in journal articles, one can say that women of the BRA show higher performance in relation to men.

Focusing on differences between academic fields, in Table 4, as expected, mean number of journal articles is higher in biological, health sciences and in exact sciences than in engineering. This difference could partially account for the women's higher general performance in the BRI, previously mentioned. But even considering journal publication in this specific group, it can also be observed that women in the biological and health areas publish, in average, more journal articles than men. Men, on the other hand, show higher performance in journal articles in exact sciences and engineering. These gender tendencies are not clear in the other two types of publication.

Table 4. Mean of types of publications of professor-researchers from Brazilian Research Institutes by sex considering academic area, Graduate Program evaluation and PHD period and PHD country – 2009.

	Publication Means					
	Journal Article		Annals Full Article		Annals Abstract	
	Women	Men	Women	Men	Women	Men
<i>GENERAL MEAN PERFORMANCE</i>	2.51	2.12	0.74	1.48	1.14	0.75
<i>ACADEMIC AREA</i>						
Engineering	0.99	1.11	2.66	2.96	0.45	0.32
Exact Sciences	2.24	2.71	1.88	1.42	0.86	0.65
Health Sciences	2.99	2.90	0.28	0.23	1.26	1.51
Biological Sciences	3.27	3.19	0.09	0.07	1.25	1.26
<i>GRADUATE PROGRAMS</i>						
Low evaluated (2 and 3)	1.12	0.90	0.99	2.07	0.98	0.30
High evaluated (6 and 7)	3.66	2.52	1.23	2.26	0.47	0.45
<i>PHD period</i>						
Before 2000	2.97	2.40	0.72	1.60	1.07	0.76
2000 and After	1.88	1.57	0.77	1.25	1.23	0.74
<i>PHD country</i>						
Brazil	2.59	2.08	0.72	1.27	1.25	0.87
Abroad	2.19	2.25	0.88	2.07	0.59	0.49

Table 4 also shows that belonging to programs with higher grades seems to have a positive impact in the output of men and women in journal articles and annals full articles. However, what stands out in the comparison of the two types of program (low and high performance) is

that women's mean number of journal articles is much higher than men's in high performance programs, where men are predominant (Table 2).

Data also suggest that professional experience, estimated through the time elapsed since PHD conclusion, contributes positively, for both women and men, to a greater output in journal publishing. On the other hand, both gender groups with more recent PHD degrees tend to publish more annals full articles. The other factor considered - PHD country- suggests that being educated abroad is more relevant to male output: men educated abroad show a much higher performance than women in this category. Regarding this last result, it could be pointed out that full articles in annals is the type of output that appears more often in the technological areas, like engineering, where 20% of the professor-researchers of the BRI are institutionally related (Table 2). It is also possible to consider that this kind of publication, which is associated to the participation in events, especially international events, may contribute to the development of professional contacts, favored by the period of experience abroad. If this is the case, women are not profiting, as much as their male colleagues, of their experience abroad.

Professor-researchers have several assignments besides publishing results based on their research projects. These assignments comprise, among others, graduate teaching, dissertation advising, banking participation and tasks involved in project leadership. How the involvement with these assignments is related with their publication output, and how gender might interfere in this process is explored in table 5.

Table 5. Mean number of involvement in academic tasks of professor-researchers from Brazilian Research Institutes by publication level and gender – 2009.

Academic Task	Professor-researchers			
	with no		with 3 or more	
	journal article		journal articles	
	Mean		Mean	
	Woman	Man	Woman	Man
Graduate Teaching	0.90	1.10	1.17	1.08
MS Advisor	0.59	0.70	0.83	0.98
PHD Advisor	0.37	0.63	0.80	0.98
Banking participation	0.94	1.42	1.00	1.18
Project Leader	0.87	0.82	1.64	1.37

Table 5 show that, in average, those BRI professor-researchers who have not published in 2009 – those with zero articles – tend to have less involvement with the different academic tasks considered, notably involvement with doctoral degree advising and project leadership. Besides, the comparison between men and women shows that men, independently of publication quantity, tend to be more involved in academic tasks, except in graduate teaching and project leadership, in which women show higher performance, but only a small positive difference. Women higher involvement in this specific task - project leadership -, especially among the more productive ones, might contribute to explain their higher performance in journal articles as previously shown in tables 3 and 4.

Concluding remarks

This work focused on gender differences in scientific production of professor-researchers attached to in BRI, aiming at identifying how institutional and background aspects may be related do their production, as well as how the diverse academic tasks performance by these men and women might interfere with their scientific production.

Considering institutional and background aspects, the results show that these professor-

researchers are allocated in the public sector, are concentrated in four academic areas, and the majority in programs that received high grades from government evaluation process (Table 1). Results also show that women are a minority in those institutes and are concentrated in the health and biological science, whereas men are concentrated in engineering and exact sciences (Table 2). Women also show higher scientific production, especially in journal articles, the most valued type of academic publication (Tables 3 and 4). Women's performance is especially outstanding when they are involved in highly evaluated graduate programs. Female professor-researchers only show lower production output in relation to their male colleagues in journal articles of traditionally masculine areas: exact sciences and engineering. But male predominance in these areas is not consistently maintained when the other types of scientific productions are considered. The last results highlighted here refer to the involvement in academic tasks by level of production. Data show that the involvement of both men and women in those tasks seems to be positively related to their productive levels, especially PHD advising and project leadership. Men, however, tend to be more involved in most academic tasks, regardless of their productive levels, with the exception of project leadership, in which women are more involved, notably the highly productive ones (Table 5). The originality of the data presented in this study is the inclusion of different types of scientific production in the analyses of gender differences in science, as well as the examination of associations of these different types of productions with contextual and academic background, as well as with involvement in academic tasks. The originality of this study is also in the selection of a particular study field: the research institutes that have an outstanding place in the development of modern science, as institutions created with the specific purpose of scientific development. Despite their relevance for the scientific field, only few studies about gender and science focus on these institutions. In Brazil, the great majority of BRI are supported by the Federal Government, are dedicated to specific scientific areas and the graduate programs under their responsibility are well recognized by the scientific community and, as data analyses shown here, tend to receive high grade marks from the national graduate programs evaluation. These indicators of excellence make it valuable the analysis of gender differences in those institutions aiming at contributing to better understand women's participation in Brazilian science and also contribute to gender governmental policy.

Intended further analyses with the BRI data will make use of statistical multivariate models trying to evaluate the relative contribution of the different contextual, background and academic tasks involvement, as well as gender in scientific production of professor-researchers. These analyses will help to indicate the importance of institutional and gender cultures, and patterns of academic practices in scientific production.

Acknowledgment

The authors are grateful to CNPq for financial support.

References

- Barrios, M., Villarroya, A., Ollé, C., & Ortega, L. (2013). Gender inequality in scientific production. *Proceedings of ISSI 2013, 1*, 811-818.
- Bonaccorsi, A. & Daraio, C. (2003). Age effects in scientific productivity. The case of the Italian National Research Council (CNR). *Scientometrics*, 58(1), 49-90.
- Bourdieu, P. (2003). *Os usos sociais da ciência*. São Paulo: UNESP.
- Brito Ribeiro, L.M.B. (2011). Gênero e Ciência: A Presença Feminina em Institutos Públicos de Pesquisa. Anais ANPAD - XXXV Encontro da Associação Nacional de Pós-Graduação em Administração –, Rio de Janeiro.
- CAPES, Cadernos de Avaliação. Received December, 2014 from <http://conteudoweb.capes.gov.br/conteudoweb/CadernoAvaliacaoServlet>.

- CAPES, GeoCapes. Distribuição de docentes, ano 2009. Received December, 2014 from <http://geocapes.capes.gov.br/geocapes2/>.
- CNPq, Areas do Conhecimento. Available at: <http://www.memoria.cnpq.br/areasconhecimento/index.htm> Accessed in December 2013.
- Epstein, C. (2007). Great divides: the cultural, cognitive, and social bases of the global subordination of women. *American Sociological Review*, 12(1), 1-25.
- European Commission. *She figures 2009: statistics and indicators on gender equality in science*. http://ec.europa.eu/research/sciencesociety/document_library/pdf_06/she_figures_2009_en.pdf.
- Fox, M. F. (2001). Women, science and academia: graduate education and career. *Gender and society*, 15(5), 654-666.
- Fox, M.F. (2010). Women and men faculty in academic science and engineering social-organizational indicators and implications. *American Behavioral Scientist*, 53(7), 997-1012.
- Gauche, S., Verdinelli, M.A., & Silveira, A. (2013). Composição das equipes de gestão nas universidades públicas brasileiras: segregação de gênero horizontal e/ou vertical e presença de homosociabilidade. Anais do IV Encontro de Gestão de Pessoas e Relações de Trabalho. Brasília, DF.
- Gillooly, P. (2011). MIT News. New report details status of women in science and engineering at MIT. Available at: <http://newsoffice.mit.edu/2011/women-mit-report-0321>
- Hermann, C. & Cyrot-Lackmann, F. (2002). Women in Science in France. *Science in Context*, 15(4), 529-556.
- INEP. (2007). A mulher na educação superior brasileira: 1991-2005 / Organizadores: Dilvo Ristoff ... [et al.]. – Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. 292 p. ISBN 85-86260-82-7.
- Leta, J. (2012). Brazilian growth in the mainstream science: The role of human resources and national journals. *Journal of Scientometric Research*, 1, p. 44-52.
- Leta, J., Olinto, G., Batista, P.D., & Borges, E.P. (2013). Gender and academic roles in graduate programs: analyses of Brazilian government data. *Proceedings of ISSI 2013*, 1, 796-810.
- Leta, J., Thijs, B., & Glänzel, W. (2013). A macro-level study of science in Brazil: seven years later. *Encontros Bibli*, 18, 51-66.
- Long, J. S. (1992). Measures of sex differences in scientific productivity. *Social Forces*, 71(1), 159-178.
- MCTI. Indicadores Ciência, Tecnologia e Inovação de 2014. Tables 3.1.2, 3.5.2, 3.5.5 and 3.6.1. Received from <http://www.mct.gov.br/index.php/content/view/740.html?execview=>
- Meulders, D., Plasman, R., Rigo, A., & O'Dorchai, S. (2010). Horizontal and vertical segregation. Meta-analysis of gender and science research – Topic report. *7th RTD Framework Programme of the European Union*.
- Mugnaini, R. & Sales, D.P. (2011). Mapeamento do uso de índices de citação e indicadores bibliométricos na avaliação da produção científica brasileira. Anais ENANCIB - Encontro Nacional de Pesquisa em Ciência da Informação. Brasília: *Thesaurus*, 12, 2361-2372.
- Olinto, G. (2011). A inclusão das mulheres nas carreiras de ciência e tecnologia no Brasil. *Inc. Soc.*, 5(1), 68-77.
- Prpic, K. (2002). Gender and productivity differentials in science. *Scientometrics*, 55(1), 27-58.
- Regalado A. (2010). Brazilian Science: Riding a Gusher. *Science*, 330(6009), 1306-1312.
- Rodrigues, J.G. (2014). A trajetória feminina na pesquisa na Fundação Oswaldo Cruz: um estudo exploratório. Tese (Doutorado em Informação, Comunicação em Saúde) – Fundação Oswaldo Cruz, Instituto de Informação Científica e tecnológica em Saúde, Rio de Janeiro, 2014.
- Shienbinger, L. (2001). O feminismo mudou a ciência? Bauru, SP: EDUSC. p. 384.
- Velema, T. (2012). The contingent nature of brain gain and brain circulation: their foreign context and the impact of return scientists on the scientific community in their country of origin. *Scientometrics*, 93(3), 893-913.
- Weltman, W. L. (2002). A produção científica publicada pelo Instituto Oswaldo Cruz no período 1900-17: um estudo exploratório. *História, Ciências, Saúde Manguinhos*, 9(1), 159-86.

Comparing the Disciplinary Profiles of National and Regional Research Systems by Extensive and Intensive Measures

Irene Bongioanni¹, Cinzia Daraio², Henk F. Moed² and Giancarlo Ruocco^{1,3}

irene.bongioanni@gmail.com, Giancarlo.Ruocco@roma1.infn.it

¹Sapienza University of Rome, Department of Physics, Rome, (Italy)

daraio@dis.uniroma1.it, henk.moed@uniroma1.it

²Sapienza University of Rome, Department of Computer Control and Management Engineering Antonio Ruberti, Via Ariosto, 25, Rome, 00185 (Italy)

³Center for Life NanoScience@LaSapienza, IIT, Sapienza University of Rome, Viale Regina Elena 295, Rome (Italy)

Abstract

In this paper, by modeling national and regional research systems as complex systems, we compare the dynamics of their disciplinary profiles using extensive (size dependent) indicators as well as intensive (size independent) average productivity indicators of scientific production. Our preliminary findings show that the differences between the disciplinary profiles among countries in the world is of the same order of magnitude of the differences among European countries, that in turn, is of the same order of magnitude of the dynamics among regions within a country. While additional research (that is in progress) is needed to confirm these findings, we describe the main advantages (features) of our approach and outline its usefulness to support evidence-based policy making.

Conference Topics

Methods and techniques; Citation and co-citation analysis; Indicators; Science policy and research assessment; Country-level studies

Introduction, scope and structure of this paper

The dynamics of national or regional research systems is one of the most important topics in quantitative science and technology research. Interestingly, a lot of studies have analyzed the disciplinary specialization of countries (see e.g. Glanzel, 2000; Glanzel & Schlemmer, 2007; Glanzel et al., 2006, 2008; Hu & Rousseau, 2009; Tian et al., 2008; Wong, 2013; Wong et al., 2012; Yang et al., 2012; Horlings & Van den Besselaar, 2013; Radosevic & Yoruk, 2014) or have investigated the disciplinary specialization of regions within a particular country, or have conducted case studies on individual regions and/or on a few number of selected disciplines (see e.g. Zhu et al., 2009; Glanzel, Tang & Shapira, 2011).

Much less studied are the disciplinary profiles of European countries at the regional level. To the best of our knowledge there are not empirical analyses at European level, investigating the evolution of the disciplinary composition (i.e. the 27 Scopus Subject categories) of regions. Moreover, none of the existing studies have analyzed in a comparative way, the range of variability (briefly: the dynamics) of national and regional research systems which is the aim of our paper. We investigate here this dynamics in terms of both extensive measures of scientific production (i.e. total number of scientific publications, citations and so on) and in terms of intensive average scientific productivity (i.e. number of publications per author).

In particular, the investigation of the dynamics of intensive measures of scientific production has an important policy relevance. According to the macroeconomic theory, we have growth convergence when smaller (poorer) countries, in terms of output per capita (e.g. GDP per capita), grow faster than larger (richer) countries. In the context of research systems, we can say that there is a convergence if smaller scientific systems, in terms of scientific output per capita, grow faster than larger one. This is an important question, related to the policy

decision of supporting catching up countries depending on whether there is convergence or not. This question is extremely important also at the regional level, for which there is an increasing interest in the smart specialization of regions, defined in terms of technological specialization, linked to the degree of innovativeness of the regions, to develop effective policies of cohesion (McCann & Ortega-Argilés, 2013; Camagni & Capello, 2013). Despite the fact that scientific specialization is commonly considered as a relevant factor for the technological specialization of regions, there is not available evidence on the scientific specialization of regions and their dynamics. Even more scant is the empirical evidence aiming at analyzing the dynamics of the scientific profiles of regions together with those at the national level, to derive informative policies to support research at national and regional level, able to take into account the complementarity/substitution relationship between national and regional research systems. We try to fill this gap, providing an investigation of the dynamics of the disciplinary profiles at the national and regional level using extensive and intensive measures.¹

Bongioanni, Daraio, Moed and Ruocco (2014) provided a first exploration at the world country level. In the current paper, the analyses are extended systematically in the following three manners.

- a) The paper analyzes a series of both extensive (size dependent) and intensive (size independent) bibliometric indicators of research productivity, impact and collaboration. Table 2 gives a list of all indicators included in the study. Data was extracted from the Scopus database and relate to the scientific production of world countries and 27 Scopus subject categories from 1996 to 2012.
- b) The analyses do not only relate to *national* research systems, but also to *regions* within European countries. In terms of the Nomenclature of Territorial Units for Statistics, NUTS-2 units were analyzed.
- c) We describe the main features and advantages of our approach to investigate the scientific convergence of national and regional research systems.

The model

A spin glass is a disordered assembly of spins (e.g. dipole magnets) that are not aligned in a regular pattern. The term “glass” comes from an analogy between the “magnetic” disorder in a spin glass and the positional disorder of a conventional, chemical glass, e.g., a window glass. In window glass or any amorphous solid the atomic bond structure is highly irregular; in contrast, a crystal has a uniform pattern of atomic bonds. In ferromagnetic solid, magnetic spins all align in the same direction; this would be analogous to a crystal. The individual interactions in a spin glass are a mixture of roughly equal numbers of ferromagnetic bonds (where neighbors prefer to have the same orientation) and antiferromagnetic bonds (where neighbors tend to orientate in the opposite directions). These patterns of aligned and misaligned magnets create what are known as frustrated interactions - distortions in the geometry of atomic bonds compared to what would be seen in a regular, fully aligned solid. They may also create situations where more than one arrangement of spins is stable.

In the physics of complex systems, a mathematical framework is developed to analyze spin glass systems. This paper uses certain elements of this framework. National or regional research systems are conceived as analoga of spins and their complex interactions give rise to disordered, spin glass like, systems. Their orientation is described in terms of the distribution of a research system’s publication output or related bibliometric measures over the various research disciplines. A research system’s disciplinary orientation is described as a vector the

¹ This is the first step of our analysis. Further research will be subsequently devoted to the exploration and investigation of the link between scientific and technological profiles of regional and national research systems.

elements of which contain the percentage of publications in the various disciplines. The rationale for using the spin glass model lies in the ability to analyze the dynamical interactions among research units in a wider system analogously to the analysis of spin orientations in spin glasses.

The following Table 1 summarizes the analogy between the main physical notions of a spin glass model and the corresponding notions in the research system model (see also the Appendix of Bongioanni, Daraio & Ruocco, 2014).

Table 1. Spin glass model: main physical notions and their corresponding notions for research system.

Notion in the physical system	Notion in the Research system
Spin	Country/region
Spin components	Scientific disciplines
J couplings	Country-to-country or region-to-region interactions
Energy (it has to be minimized to find stable solutions)	Generalized cost function (to be minimized)
Overlap	Similarity measure

Within the framework of this model, Bongioanni, Daraio & Ruocco (2014) proposed to compare the disciplinary patterns of research systems, by computing the ‘overlaps’ quantities, that are similarity measures between disciplinary patterns, borrowed from the physics of complex systems. The main variables analysed here are the $P_a(i)$ i.e. the shares of articles published in a subject category i for a given country (or region) a over the sum of publications made during 1996-2012. Similar variables are based on the number of citations received, or the number of internationally co-authored papers. Table 2 gives an overview of all indicators used in this study. The measure of the overlap between the pattern of disciplinary profiles of two countries a and b , $P_a(i)$ and $P_b(i)$ respectively, that is the measure of similarity between systems, is defined as:

$$q_{ab} = \frac{1}{D} \sum_{i=1}^D \sigma_a(i) \sigma_b(i),$$

where

$$\sigma_a(i) = \frac{P_a(i) - \langle P_a \rangle}{\sqrt{\langle P_a^2 \rangle - \langle P_a \rangle^2}},$$

in which $\langle A \rangle$ stands for average of A , $\sigma_a(i)$ and $\sigma_b(i)$ represent the normalised shares of the indicator considered, for country (or region) a and b , respectively; and D is the number of subjects or disciplines analysed, which in this study amounts to 27 and are derived from Scopus. We note that if we use as variables $\pi_a(i) = P_a(i) - \langle P_a \rangle$ instead of $P_a(i)$, q_{ab} coincides with the Salton’s cosine (calculated with the variables π).

The overlap measure or similarity of profiles between two countries a and b , q_{ab} , ranges from -1 , meaning precisely the opposite profile, to 1 , meaning precisely the same profile, with 0 representing independence and intermediate values indicating in-between levels of similarity or dissimilarity. Moreover, the overlap can be calculated with respect to another country, with respect to an average or standard value or with respect to a given distribution.

Interpreting the distribution of the overlaps to shed lights on the dynamics of the overall system.

An interesting property of the computed overlap measures between two countries (or regions)' profiles relates to their distribution. The distribution of the overlap reveals whether there is a *convergence* in the overall system towards a unique disciplinary profile or whether there is a divergence of the system towards different disciplinary configurations. In particular, according to Bongioanni, Daraio and Ruocco (2014) the interpretation of the distribution of the overlap values is as follows: one pick on one shows a convergence towards the *same* disciplinary profile for all countries, while two picks point to two *different* configurations of disciplinary profiles.

We point out that this is one of the main advantages of our approach compared to currently bibliometric approaches used for comparing disciplinary profiles. Although a systematic comparison of our approach with other existing methods is in progress, we think that our approach offers an easy way, based on the investigation of the distribution of the overlap, to check whether there is convergence or not without having to adopt one of the alternative methods developed in the theory of growth to measure convergence. The most applied method to assess convergence in this context, adopted also in the context of scientific convergence (see e.g. Horlings & van den Besselaar, 2013), is based on *regressions*. Within this framework (see e.g. Barro & Sala-i-Martin, 1992), it is said that there is *beta-convergence* (where beta is the coefficient of the initial level of per capita output in the growth regression) when poor economies tend to grow faster than rich economies (and hence the beta coefficient is lower than zero, implying that the higher initial level of output per capita negatively affects the growth rate). Another related concept is that of *sigma-convergence*, which happens when the dispersion of the output per capita decreases over time. The sigma-convergence is often measured by analyzing the variation of the standard deviation (or the coefficient of variation or the concentration) of the output per capita over time. However, this regression based approach has been questioned in the growth literature (see e.g. Durlauf, 2000) and other studies of convergence have applied different methods, including a test on the distribution of the output and how it evolves over time, reaching often very different results (see e.g. Durlauf, Kourtellos, & Tan, 2005). Our approach, offers an interesting alternative to estimate the convergence, by analyzing the distribution of the overlaps and their dispersion.

Another interesting property of our approach is related to the exploitation of the *ultrametric* structure of the overlap values to obtain “automatically” clusters of the national or regional research systems analysed, without having to carry out a specific clustering exercise.²

Note that the indicators reported in bold in Table 2 are average productivity indicators, that is intensive (size independent) indicators of the scientific production, while the others are extensive (size dependent) indicators of scientific production.

In this paper the following overlaps were computed:

- Of each main country in the world against all other countries, using a set of 41 countries, including all member states of the European Union and major countries from the rest of the world.
- Of each 27 European country against all other European countries, to provide an aggregate benchmark for the regional analysis.
- Of each NUTS-2 region against all other regions, using a set of 266 NUTS-2 regions in member states of the European Union.

² Research on this point is in progress.

Table 2. Indicators applied in the study

Indicator	Description
PUB	Number of articles (integer count).
PUBf	Number of articles (fractional counts based on authors affiliations).
C	Total citations (4 years window, i.e., for articles in 2006; citations are from 2006-2009).
CPP	Total citations per paper (4 years window, i.e., for articles in 2006; citations from 2006-2009).
HCPUB	Number of articles in top 10 per cent of most highly cited articles in a discipline.
PUBINT	Number of internationally co-authored papers.
PUBNAT	Number of nationally (but not internationally) co-authored papers.
PUBINST	Number of papers co-authored by members of different institutions within a country.
PUBSA	Number of non-collaborative (single address) papers.
NA	Number of publishing authors in a particular year, by discipline.
APUB	Number of articles (integer count) divided by NA
APUBf	Number of articles (fractional counts based on authors affiliations) divided by NA
AC	Total citations (4 years window, i.e., for articles in 2006; citations are from 2006-2009) divided by NA
ACPP	Total citations per paper (4 years window, i.e., for articles in 2006; citations from 2006-2009) divided by NA
AHCPUB	Number of articles in top 10 per cent of most highly cited articles in a discipline divided by NA
APUBINT	Number of internationally co-authored papers divided by NA
APUBNAT	Number of nationally (but not internationally) co-authored papers divided by NA
APUBINST	Number of papers co-authored by members of different institutions within a country divided by NA
APUBSA	Number of non-collaborative (single address) papers divided by NA

Legend to Table 2: Data was extracted from the Scopus database and relate to the scientific production of world countries and NUTS2 European regions for 27 Scopus subject categories from 1996 to 2012.

Results are presented in two sections. The first part explains the base notion of a disciplinary profile, compares pair-wise profiles of countries and NUTS2 regions, and analyzes the structure within the set of profiles. It focuses on one single indicator: the number of articles (PUBf) published in 2012. The second part analyzes also average productivity indicators (APUBf) and dynamical aspects.

Disciplinary profiles of countries and regions

Figure 1 shows large differences in the distribution of research articles among subject fields between USA and China. The first country has a strong focus on medical sciences and biomedical research, including biochemistry, genetics and molecular biology, neurosciences, and on social sciences and humanities. The latter shows a large publication activity in physical sciences and engineering: chemistry, materials science, physics, and engineering and computer science.

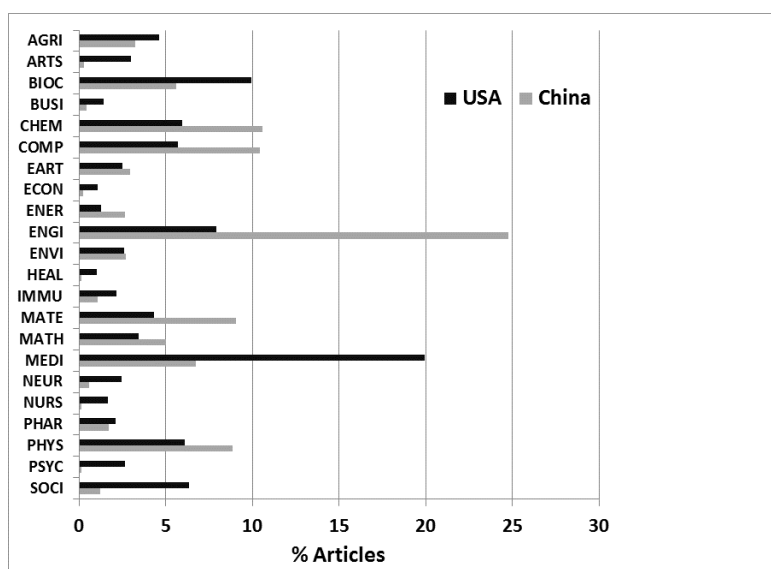


Figure 1. Disciplinary profiles of two countries large countries: China vs. USA. Data relate to the year 2012, and are extracted from Scopus.³ In this figure, four small disciplines have been left out: Dentistry, Decision Sciences, General, and Veterinary Sciences. Chemical Engineering is merged with Chemistry.

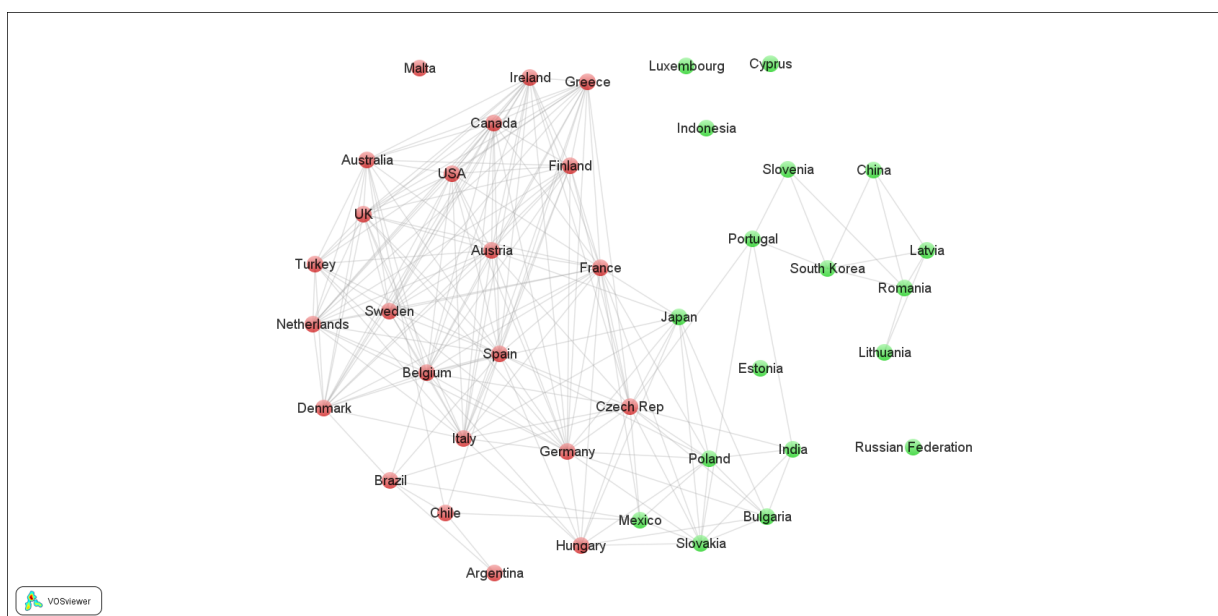


Figure 2. VoS-Viewer Map of the de degree of overlap of disciplinary profiles among 41 countries. For more details about VoS viewer, the reader is referred to www.vosviewer.com

Figure 2 shows a map of a set of 41 countries, including all member states of the European Community, and major countries from the rest of the world. Interestingly, the cluster module in the VoS Viewer identified two clusters of countries. These clusters correspond to the

³ The labels of the disciplines are the following: AGRI: Agricultural and Biological Sciences; ARTS: Arts and Humanities; BIOC: Biochemistry, Genet, Mol Biol; BUSI: Business, Managmnt, Accounting; CHEM: Chemistry; COMP: Computer Science; DECI: Decision Sciences; DENT: Dentistry; EART: Earth and Planetary Sciences; ECON: Economics, Econometrics and Finance; ENER: Energy; ENGI: Engineering; ENVI: Environmental Science; GENE: General; HEAL: Health Professions; IMMU: Immunology and Microbiology; MATE: Materials Science; MATH: Mathematics; MEDI: Medicine; NEUR: Neuroscience; NURS: Nursing; PHAR: Pharmacology, Toxicology and Pharmaceutics; PHYS: Physics and Astronomy; PSYC: Psychology; SOCI: Social Sciences; VETE: Veterinary Sci.

different profiles illustrated in Figure 1. The countries indicated with red circles, located at the left hand side of the plot, tend to have a biomedical disciplinary profile, similar to USA and the Netherlands. At the right hand side a group of countries indicated by green circles tends to have a physical-sciences profile, like China, and Russia. Many Central and Eastern-European countries belong to this group: apart from South Korea, also India, Indonesia, Mexico, Portugal, and the small countries Luxembourg and Cyprus.

Several studies in the past have found differences in disciplinary profiles between countries. But to the best of our knowledge, no study has systematically analyzed geographical regions within countries. Figures 3 and 4 show results for the so called NUTS-2 regions. In total, 266 NUTS2 regions were identified. Table 3 presents the quantiles of the distribution of the number of published articles (year 2012) among regions. The distribution is highly skewed. The top 25 per cent of regions has published more than 4,146 articles in 2012. 5 per cent has published more than 11,612 articles. The bottom 25 per cent has published less than 496, and the bottom 10 per cent less than 89. Figure 3 shows disciplinary profiles of two pairs of NUTS2 regions: Inner London and the German city Stuttgart. The figure reveals the same main profiles as Figure 1 did at the level of countries: a biomedical profile in Inner London, and a physical sciences profile in Stuttgart.

Table 3. Quantiles of the distribution of number of publications among NUTS2 regions

<i>Level</i>	<i>Score</i>
Number of NUTS2 regions	266
Average articles/region	3,326
<i>Level</i>	<i>Quantile</i>
100% Max	46,451
90%	8,247
75% Q3	4,146
50% Median	1,815
25% Q1	496
10%	89
0% Min	1

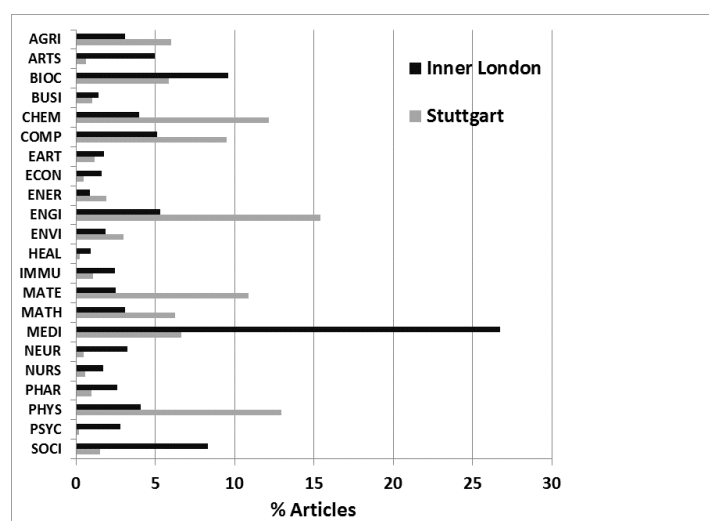


Figure 3. Disciplinary profiles of Inner London (UK) vs. Stuttgart (Germany)

different profiles illustrated in Figure 1. The countries indicated with red circles, located at the left hand side of the plot, tend to have a biomedical disciplinary profile, similar to USA and the Netherlands. At the right hand side a group of countries indicated by green circles tends to have a physical-sciences profile, like China, and Russia. Many Central and Eastern-European countries belong to this group: apart from South Korea, also India, Indonesia, Mexico, Portugal, and the small countries Luxembourg and Cyprus.

Several studies in the past have found differences in disciplinary profiles between countries. But to the best of our knowledge, no study has systematically analyzed geographical regions within countries. Figures 3 and 4 show results for the so called NUTS-2 regions. In total, 266 NUTS2 regions were identified. Table 3 presents the quantiles of the distribution of the number of published articles (year 2012) among regions. The distribution is highly skewed. The top 25 per cent of regions has published more than 4,146 articles in 2012. 5 per cent has published more than 11,612 articles. The bottom 25 per cent has published less than 496, and the bottom 10 per cent less than 89. Figure 3 shows disciplinary profiles of two pairs of NUTS2 regions: Inner London and the German city Stuttgart. The figure reveals the same main profiles as Figure 1 did at the level of countries: a biomedical profile in Inner London, and a physical sciences profile in Stuttgart.

Table 3. Quantiles of the distribution of number of publications among NUTS2 regions

<i>Level</i>	<i>Score</i>
Number of NUTS2 regions	266
Average articles/region	3,326
<i>Level</i>	<i>Quantile</i>
100% Max	46,451
90%	8,247
75% Q3	4,146
50% Median	1,815
25% Q1	496
10%	89
0% Min	1

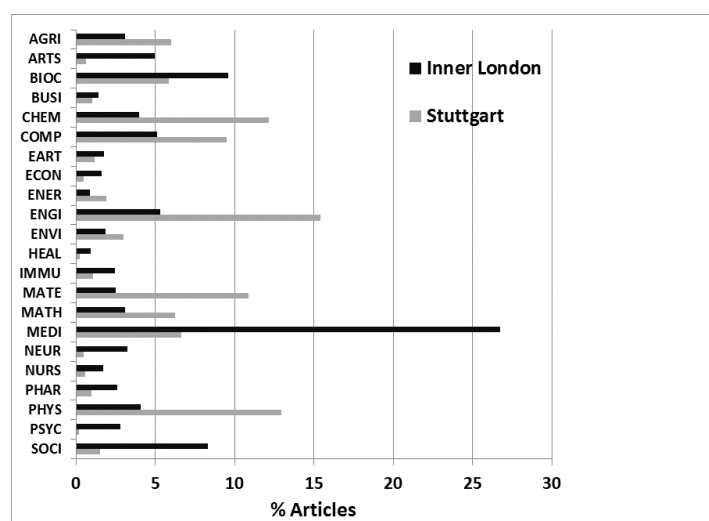


Figure 3. Disciplinary profiles of Inner London (UK) vs. Stuttgart (Germany)

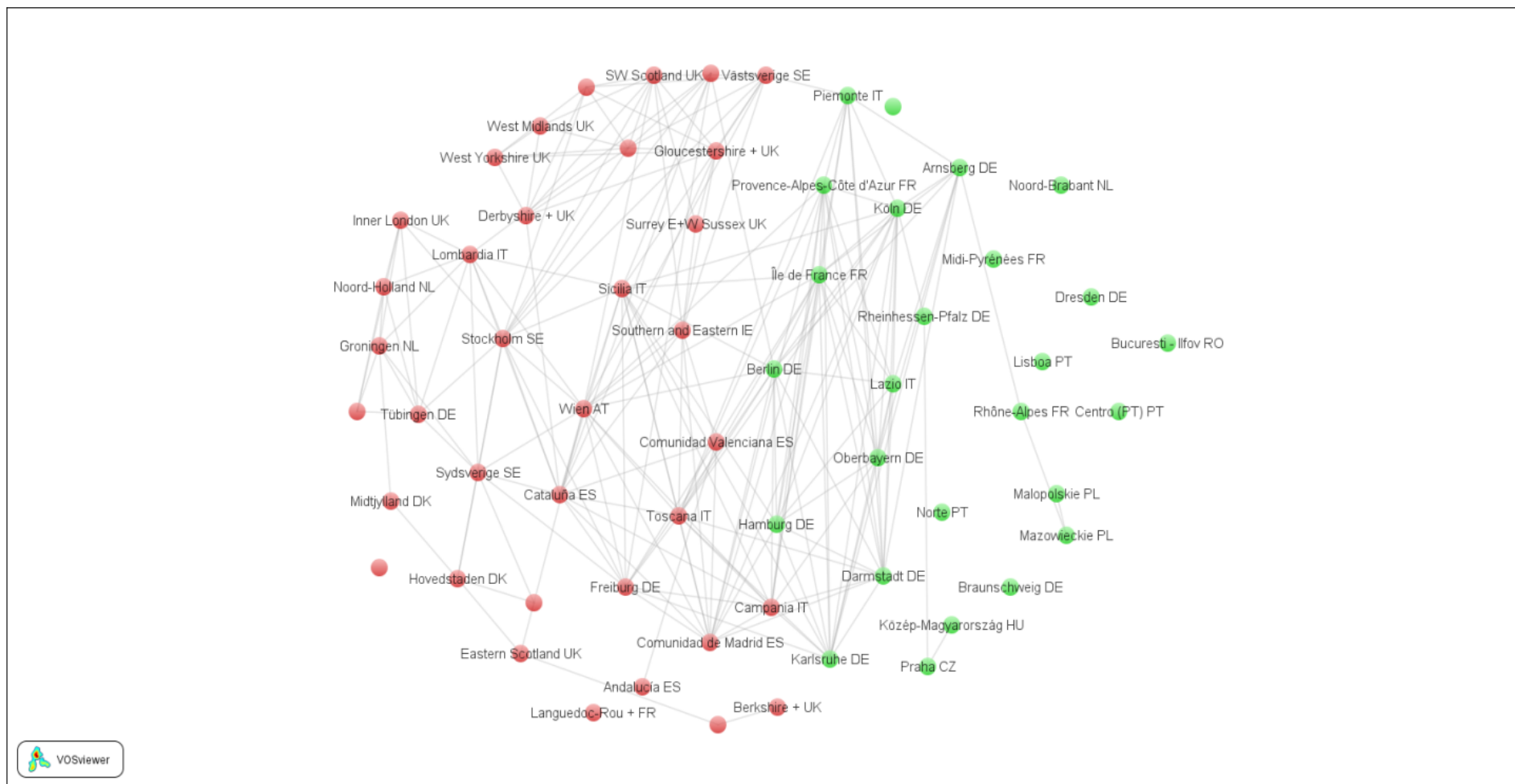


Figure 4. VoS-Viewer Map of the de degree of overlap of disciplinary profiles among 62 NUTS 2 regions. Results are based on an analysis of 62 NUTS2 regions. Due to inconsistencies in the primary data, regions from Belgium and Finland are missing in this graph. Not all circles have labels.

Figure 4 presents a VoS viewer map of the 62 NUTS2 regions in the top quartile in terms of number of articles published in 2012, and based on their degree of overlap between disciplinary specialization. As for countries, the clustering module identified two clusters: the one on the right hand side with red labels tend to cover the regions with a predominantly biomedical profile, and the cluster at the right hand side the regions with a focus on physical sciences. Due to particularities of the underlying primary data and of the visualization technique, this figure cannot be used to reliably assess regions in terms of their scientific performance. Its main function in this paper is analyzing the structure within the set of NUTS2 regions. A preliminary results that should be substantiated in further empirical analysis is that the variability of disciplinary profiles *among countries*, is of the same order of magnitude of the variability *among regions* within a country.

Analysis of distributions of overlap values

Figure 6 (see next page) illustrates the nonparametric kernel distributions (solid line) as well as the histogram of the overlap values calculated at the world, European and regional NUTS2 level. On the x-axis the overlap values are reported while on the y-axis the distribution of the overlap ($F(q)$, given by the nonparametric kernel density and the histogram) is reported. The overlaps are calculated over the volume of publications in fractional count (PUBf) as well as on the average productivity (APUBf). Remarkably, all the distributions of the overlaps clearly show a pick on one reflecting, as explained in Bongioanni, Daraio & Ruocco (2014), the existence of a *convergence towards a unique disciplinary profile*, both in extensive and intensive measures. We observe however that the distributions of the average productivity (APUBf) is *less dispersed* than that of the corresponding extensive measure at all the three levels of analysis: world, European countries and European regions. A similar pattern was found for the citation-based indicator: the number of highly cited articles published from a country or a region (HCPUB). The relative figures are not reported to save space.

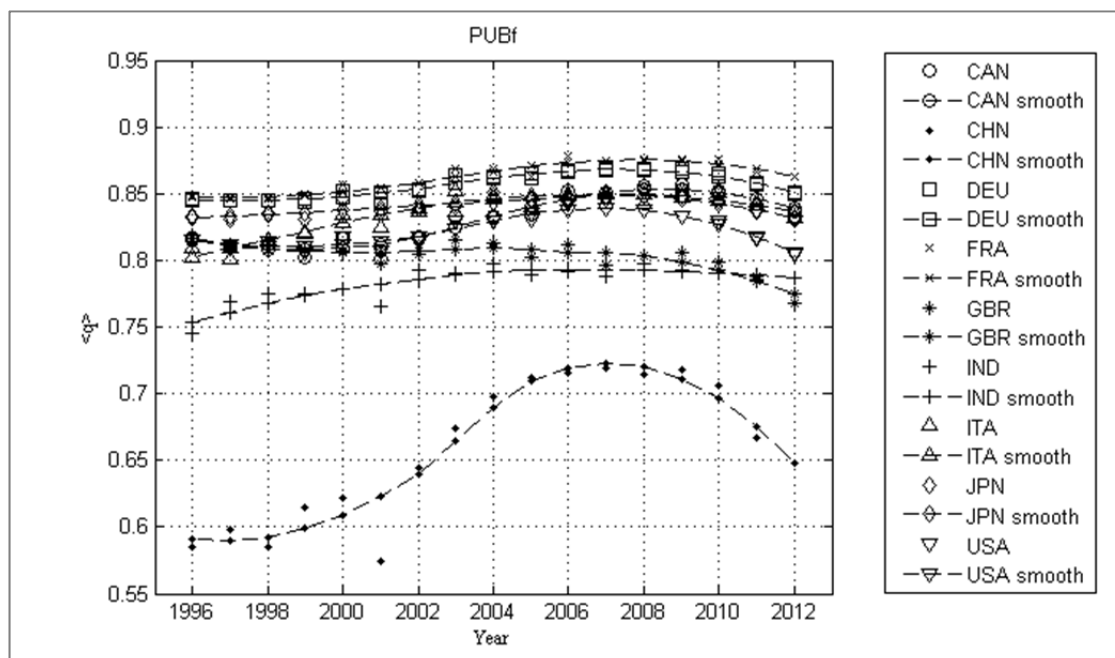
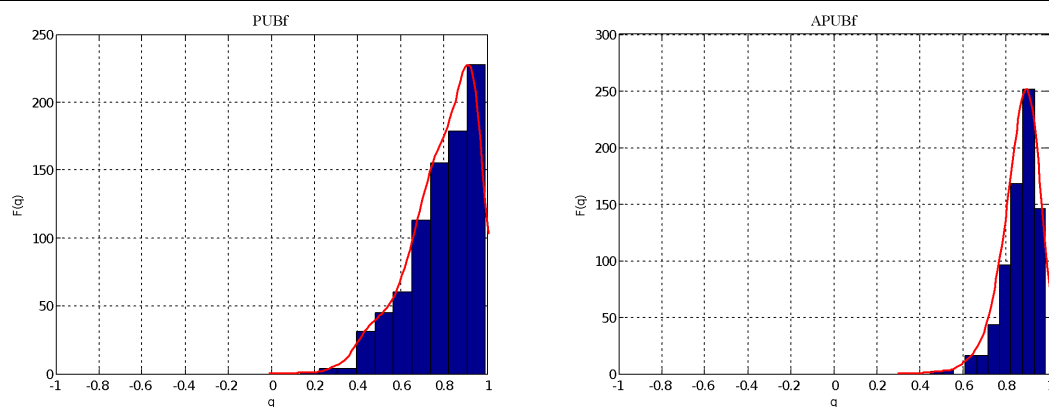
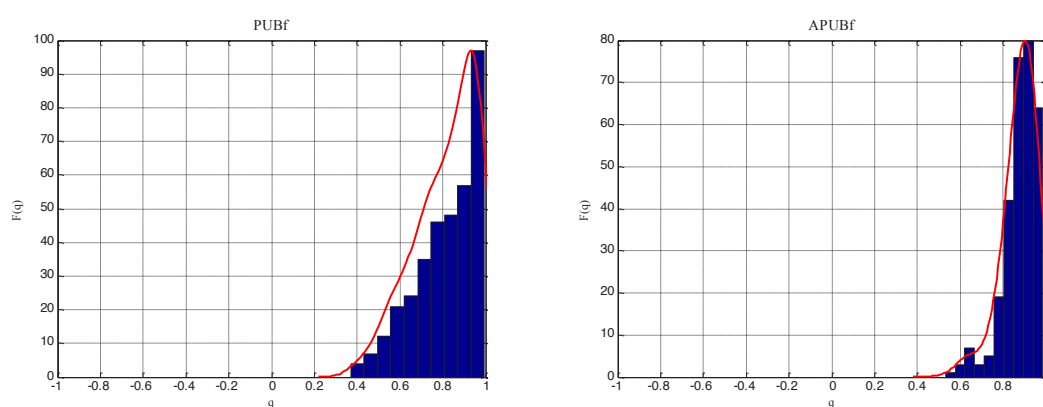


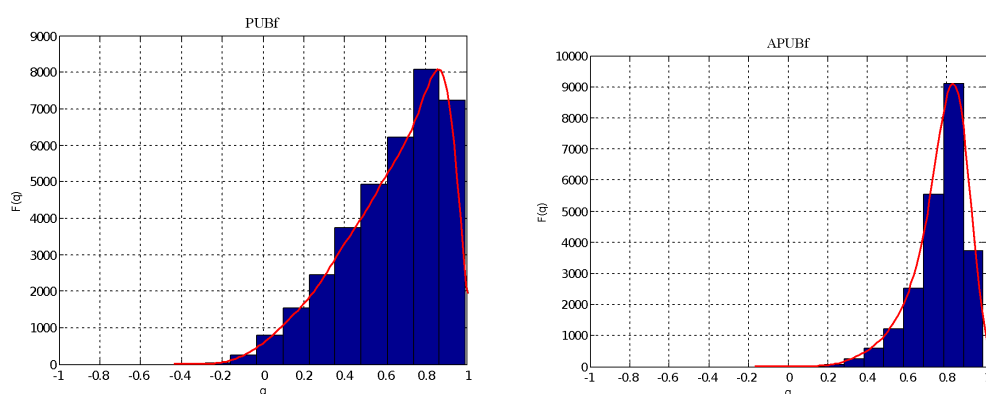
Figure 5. Dynamics of overlaps between 9 leading nations and all other countries for the fractional number of publications (PUBf).



TOP PANEL. World Distribution of the overlaps calculated on each country against all other countries in the world for the *extensive* (size dependent) indicator of scientific production PUBf (top-left panel) and the *intensive* average productivity indicator APUBf (top-right panel).



MIDDLE PANEL. European Distribution of the overlaps calculated on each European country against all other European countries for the *extensive* indicator of scientific production PUBf (middle-left panel) and the *intensive* average productivity indicator APUBf (middle-right panel).



BOTTOM PANEL. European Regions (NUTS2 units) Distribution of the overlaps calculated on each European region against all other European regions for the *extensive* indicator of scientific production PUBf (bottom-left panel) and the *intensive* average productivity indicator APUBf (bottom-right panel).

Figure 6. Distributions of the overlaps calculated at World, European and Regional level for extensive (PUBf) and intensive (APUBf) indicators.

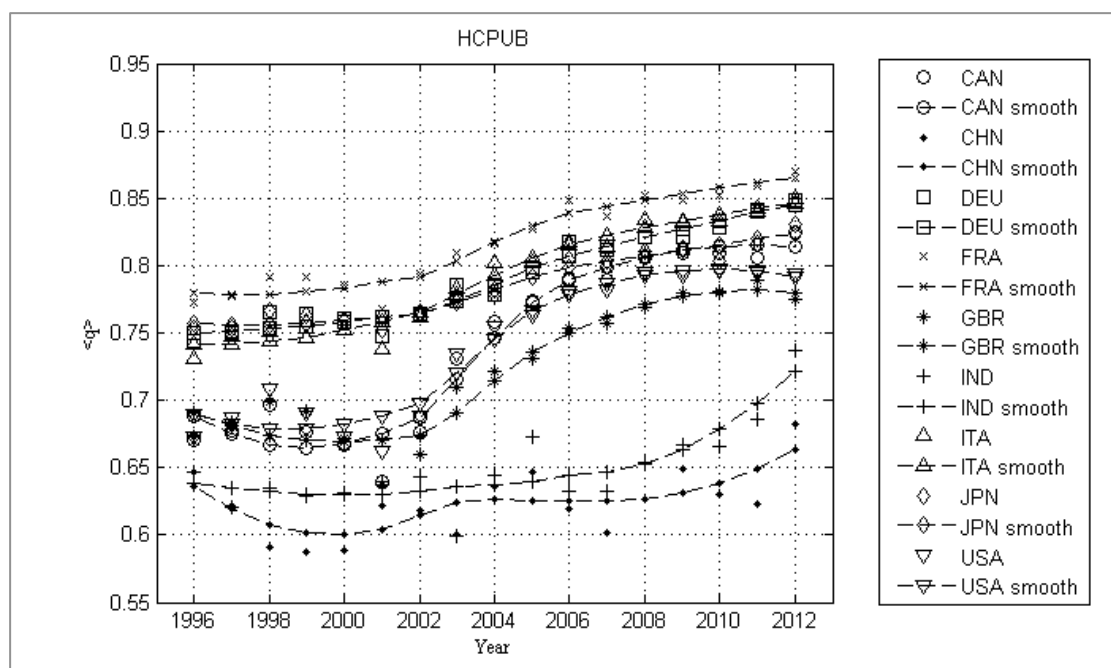


Figure 7. Dynamics of overlaps between 9 leading nations and all other countries for the number of highly cited publications (HCPUB)

An important aspect is the *dynamics* of the overlap values: how do the overlap distributions develop over time, and how does the position of specific countries evolve. Figures 5 and 7 present for 9 leading nations the development over time of the average overlap with all other countries, for the fractional number of publications (PUBf) and the number of highly cited publications (HCPUB), respectively. Although Figure 6 shows during the last 4 years a slight decline in overlap for most countries, Figure 7 reveals a trend towards convergence, especially for India and China. Perhaps the latter two countries increased their contribution to the international research front, but they maintained to some extent their own disciplinary profiles.

Conclusions

A tentative conclusion that should be substantiated in future empirical research is that the *variability of disciplinary profiles among countries is of the same order of magnitude of the variability among regions within a country and that the same happens for their convergence rates*, as shown by the distributions of the overlap calculated and displayed in this paper. The same dynamics observed for the extensive measures of scientific production is observed for the intensive average productivity, which appears to have a more concentrated distribution for all the level of the analysis carried out. Further research is in progress to support these preliminary findings and to illustrate the advantages of our approach, including the application of the ultrametric property of the overlap values to determine “automatic” clustering of the investigated national and regional systems of research. The step further will be then to link the scientific structure of national and regional systems with their technological structure to evaluate their dynamics at national and regional level.

Acknowledgments

This work was supported by Elsevier that provided the data within the Elsevier Bibliometric Research Programme (EBRP) for the Project “Assessing the Scientific Performance of Regions and Countries at Disciplinary Level by Means of Robust Nonparametric Methods: New Indicators to Measure Regional and National Scientific Competitiveness”.

References

- Barro, R. J. & Sala-i-Martin, X. (1992), Convergence, *Journal of Political Economy*, 100(2), 223–251.
- Bongioanni, I., Daraio, C., & Ruocco, G. (2014), A Quantitative Measure to Compare the Disciplinary Profiles of Research Systems and their evolution over time, *Journal of Informetrics*, 8, 710–727.
- Bongioanni, I., Daraio, C., Moed, H.F., & Ruocco, G. (2014), Disciplinary Profiles and Performance of Research Systems: a World Comparison at the Country level, *Proceedings of the Science and Technology Indicators Conference 2014 “Context Counts: Pathways to Master Big and Little Data”*, 3-5 September 2014, edited by E. Noyons, published by Universiteit Leiden CWTS 2014, pp. 50-63, ISBN 978-90-817527-1-8.
- Camagni, R., & Capello, R. (2013). Regional innovation patterns and the EU regional policy reform: Toward smart innovation policies. *Growth and Change*, 44(2), 355-389.
- Durlauf, S. (2000), “Econometric Analysis and the Study of Economic Growth: A Skeptical Perspective,” in *Macroeconomics and the Real World*, R. Backhouse and A. Salanti, eds., Oxford: Oxford University Press.
- Durlauf, S. N., Kourtellos, A., & Tan, C. M. (2005). Empirics of growth and development. *International Handbook of Development Economics*, 1.
- Glanzel, W., Debackere, K., & Meyer, M. (2008). ‘Triad’ or ‘tetrad’? On global changes in a dynamic world. *Scientometrics*, 74, 71-88.
- Glänzel, W., Leta, J., & Thijs, B. (2006). Science in Brazil. Part 1: A macro-level comparative study. *Scientometrics*, 67(1), 67-86.
- Glanzel, W. (2000). Science in Scandinavia: A bibliometric approach. *Scientometrics* 48, 121-150.
- Glanzel, W. & Schlemmer, B. (2007). National research profiles in a changing Europe (1983–2003): An exploratory study of sectoral characteristics in the triple helix. *Scientometrics*, 70(2), 267–275.
- Hu, X. J., & Rousseau, R. (2009). Comparative study of the difference in research performance in biomedical fields among selected Western and Asian countries. *Scientometrics*, 81(2), 475–491.
- Horlings, E. & van den Besselaar, P. (2013), Convergence in science growth and structure of worldwide scientific output, 1993-2008, Rathenau Instituut, Working Paper 1301.
- McCann, P. & Ortega-Argilés, R. (2013). Smart specialization, regional growth and applications to European Union cohesion policy. *Regional Studies*, 1-12.
- Radosevic, S., & Yoruk, E. (2014). Are there global shifts in the world science base? Analysing the catching up and falling behind of world regions. *Scientometrics*, 101(3), 1897-1924.
- Tang, L. & Shapira, P. (2011). Regional development and interregional collaboration in the growth of nanotechnology research in China. *Scientometrics*, 86, 299–315.
- Tian Y., Wen, C., & Hong, S. (2008). Global scientific production on GIS research by bibliometric analysis from 1997 to 2006. *Journal of Informetrics*, 2, 65-74.
- Wong, C. Y. (2013). On a path to creative destruction: science, technology and science-based technological trajectories of Japan and South Korea. *Scientometrics*, 96, 323–336.
- Wong, C. Y. & Goh, K. L. (2012). The pathway of development: science and technology of NIEs and selected Asian emerging economies. *Scientometrics*, 92, 523–548.
- Yang, L. Y., Yue, T., Ding, J. L., & Han, T. (2012). A comparison of disciplinary structure in science between the G7 and the BRIC countries by bibliometric methods. *Scientometrics*, 93, 497–516.
- Zhou, P., Thijs, B., & Glänzel, W. (2009), Regional analysis on Chinese scientific output, *Scientometrics*, 81(3), 839-857.

New Research Performance Evaluation Development and Journal Level Indices at Meso Level

Muzammil Tahira¹, Rose Alinda Alias¹, Aryati Bakri¹ and A. Abrizah²

¹*mufals@yahoo.com*, ¹*alinda@utm.com*, ¹*aryatib@utm.com*

¹Department of Information System, Faculty of Computing,
Universiti Teknologi Malaysia (UTM), Skudai 81310, Johor Bahru, (Malaysia)

²*abrizah@um.edu.my*

²Department of Library & Information Science,
Faculty of Computer Science & Information Technology, University of Malaya, 50603 Kuala Lumpur,
(Malaysia)

Abstract

This study applies scientometric approach to meso level data. The objective was to evaluate Institutional level h-index's (IHI) reliability with respect to other Journal Related Indices (JRI). Most of the studies in the literature considered journal's h-index as contrasted measure. Nevertheless, there has been no study that explores the relation between IHI and institutional level JRI. To get further evidence, we have explored the inter-correlation of IHI with a set of JRI. For this purpose data from Web of Science, Journal Citation Report and time cited features were used. Our unit of analysis was Malaysian engineering research with a wider time span of 10 year's data (2001-2010) and a larger set of journals (1381 journals). Previous studies are used for comparative analysis. This paper puts forward a better understanding to considering new impact indices at meso level for evaluation purpose.

Conference Topic

University policy and institutional rankings, Science policy and research assessment

Introduction

Journal Impact Factor (JIF) was introduced by the Institute of Scientific Information (ISI) via Journal Citations Report (JCR) about 30 years ago. It has a long tradition as an Impact Factor (IF) indicator for scholarly research output. Alike, h-index and many of its variants have been introduced and displayed on JCR site (www.webofknoweldge.com). IF can be used as a measure of research quality/impact of journals (Braun, Glanzel & Schubert, 2006). In general research performance evaluation (RPE) practices, it has become a "chief quantitative measure of the quality of researcher, and even the institution" but, it cannot be used as a direct measure of quality (Amin & Mabe, 2003; Bornmann et al., 2011). JIF remains the primary criterion when it comes to assessing the quality of journals and authors (Raj & Zainab, 2012). IF should not be used as a sole measure of a journal rank (Bornmann, et al., 2011).

To overcome the limitations, of IF, researchers suggested that it should be used with new alternative tools (Braun, Glanzel, & Schubert, 2006; Prathap, 2011; Bornmann et al., 2011; Yang Yin, 2011) or as a measure of research quality / impact of journals (Braun, Glanzel & Schubert, 2006). An interesting debate was started by Braun, Glanzel, and Schubert, (2006) who suggested that the h-index can be used as a measure of research quality or impact of a journal. The notion of Journal h-index was introduced by (Braun, Glanzel, & Schubert, 2005). Who found it a promising measure for the journal (Braun, Glanzel, & Schubert, 2006). After the introduction of h-index, a number of studies made a comparative analysis of both measures and their variants. Both impact indices (h and IF) are easily comprehensible (Leydesdorff, 2009) and have received worldwide recognition. However, prior studies, as reviewed in the subsequent paragraphs were concerned with the evaluation of journal's h-index to JRI.

Mingers, Macri and Petrovici (2012) examined Journal level h-index against Impact Factor 2year (JIF), Impact Factor 5 year (IF5y) and peer judgment for management journals. They preferred journal h-index to IF because of the former's selective time frame and the formulaic problem. Another study in the field of management was carried out by Moussa and Touzani (2010) using Google-Scholar (GS) as source data. They used a variant of the h-index, the hg-index along with two and five years IF. There was a substantial agreement found (>0.85) between JIF 5y and the hg-index ranking. They suggested hg-index as an alternative to the GS based journals. Soutar and Murphy (2009) studied 40 marketing journals and ranked them according to IF and h-index, and compared their list with Australian journal ranking. They suggested these indices as the basis for moving some journals up and other journals down. Their study supported the use of GS as an alternative way to measure citations in marketing. Harzing and Van der Wal compared h-index calculated from GS with the impact factors computed from the Web of Science (WoSTM) and with peer reviewed journal ranking (2009) by undertaking a larger-scale investigation of over 800 business and management journals.

A comparative analysis of IF and h-index was carried out by Bador and Lafouge (2010) on pharmacology and psychiatry journals from JCR with two-year publications. The journals correlation coefficient between IF and h-index was high. They inferred that IF and h-index can be totally corresponding when analyzing journals of the similar scientific subject. Bornmann, Mutz and Daniel (2009) studied the journal's h-index of twenty organic chemistry journals from WoSTM database for two years time span. They analyzed a number of impact indicators including the IF, and journal's h-index and its variants g index, h^2 index, A , and R index. They found "a high degree of correlation between the various measures" (Bornmann, Mutz & Daniel, 2009).

Yang Yin (2011) analyzed 20 top journals in the field of science and engineering using data from WoSTM. The researcher hypothesized "that the combination of complementary journal indicators could provide a simple, flexible and practical alternative approach for evaluating scientific journals" (p.2). Yang Yin considered the journal h-index with another JRI e.g. EigenFactor score. There is a positive correlation although not strong among these indices. They suggested getting published research work in high Eigenfactor scores journals. These indices can also be combined to complement each other.

Research Objectives

The objective of past studies was to evaluate a journal's h-index validity and reliability with respect to other JRI. Most of these studies considered journal's h-index as contrasted measure with JIF, JIF (5Y), and EigenFactor Score (EF). These studies are meaningful to understand the properties of newly introduced indices and potential use of journal's h-index as a complement aid with IF and its variants (Bador & Lafouge, 2010; Bornmann et al., 2012; Yang Yin, 2011) or, as a supplement (Braun, Glanzel, & Schubert, 2006).

Nevertheless, there has been no study to explore the relation of IHI with JRI. To have further evidence of validity of h-index at the institutional level, we hypothesized that IHI is a potential index for RPE that can be used to complement or as a supplement along with JRI for RPE at the institution level.

Methods and Materials

The empirical part of this study focuses on one non-Western country, Malaysia, which has a developed and well-defined scholarly publishing industry based in its universities. Research productivity, citations record, and institutional journal data of twelve Malaysian universities are retrieved from WoSTM and JCR'2011 from the Web of Science. Only those universities that have at least fifty publications during the past ten years were selected. "The statistical methodology of EFA can be used to examine for latent associations present in a set of

observed variables, and reduce the dimensionality of the data to a few representative factors” (Schreiber et al., 2012, p.349). It is mainly used to identify a smaller set of salient variables from a larger set and to explore the underlying dimensions or factors that explain the correlations among a set of variables (Conway and Huffcutt, 2003). Initially, we used eleven indices for the present study. These are Total publications (TP), Total Citations (TC) Citation Per Publications (CPP), Institutional H-Index (IHI), JIF, Cumulative Journal Impact Factor (CIF), Journal Impact Factor 5y (JIF5y), Cumulative Journal Impact Factor 5y (CJIF5y), Average Impact Factor (AIF), Median Impact Factor (MIF), Immediacy-index (Imm-index) and EigenFactor Score (EF). The definitions and the acronym used are described in Table 1.

Table 1. Definitions of indices used at Meso level.

<i>Indicators</i>	<i>Definition</i>
1. Total Publications (TP)	Total publications of a university over the set criteria
2. Total Citations (TC)	Total citations of a university over the set criteria
3. Institutional H-Index (IHI)	An institution has index h if h of institutional publication has at least h citation each and other publication have fewer than or equal to h citations each.
4. Journal Impact Factor (JIF)	The average number of times articles from the journal published in the past two years has been cited in the JCR year (Thomson- Reuters 2015).
5. Cumulative Journal Impact Factor (CIF)	This is the cumulative value of Journal Impact Factor of each university.
6. Impact Factor five Years (IF5y)	The average number of times articles from the journal published in the past five years have been cited in the JCR year (Thomson-Reuters 2015).
7. Cumulative Impact Factor Five Years (CIF5y).	This is the cumulative value of five years Journal Impact Factor of each university.
8. Average Impact Factor (AIF)	This is the average of the Impact Factor of each university.
9. Median Impact Factor (MIF)	This is the median of the Impact Factor of each university.
10. Immediacy-index (Imm-index)	This is calculated by dividing the number of citations to articles published in a given year by the number of articles published in that year Thomson-Reuters 2015).
11. EigenFactor Score(EF)	“Eigenfactor score is calculated by the ratio of the total number of citations for the JCR year to the total number of articles published in the last 5 years”. Thomson-Reuters 2015).

Data Processing

To get a meaningful evaluation, we used a wider set of WoSTM engineering journals (1381 journals) considered by our sample (12 Malaysian universities) institutions with a wider horizon of ten years (2001-2010) under specified nine categories. Our research term was “Malaysia” in “Address”, limited to document type research article and reviews only and

refined by nine engineering research categories. These engineering categories are engineering electrical, electronic, engineering manufacturing, engineering biomedical, engineering industrial, engineering civil, engineering chemical, engineering mechanical, engineering environmental and engineering multidisciplinary.

Data were suffered from affiliation problem, change of journal title and abbreviation of a journal name. All the data were checked manually for publications, citations, institutional affiliation, and journal name change or emergence cases. The selected twelve universities got their articles published in 1381 journals. According to JCR'2011, almost all journals in our data set were IF. There were only 22 journal articles published in six journals, and ten proceedings had no impact factor. It is assumed that the said journals/proceedings may have IF prior to 2011. These records were included in the journal list for analysis purpose. Firstly, all the records were retrieved in a spreadsheet file, and IBM SPSS version'19 was used for statistical analysis purpose.

Table 2 provides the university-wise total journal records. The publication share of research university (RU) status was 66 % (908) while; the non-RU status universities shared 34 % (473) of the total journals.

Table 2. Distribution of journals (N=1381).

<i>No</i>	<i>University</i>	<i>Total journals and proceedings</i>	<i>University Status</i>	<i>Contribution%</i>
1	University of Malaya (UM)	191		
2	Universiti Sains Malaysia (USM)	188		
3	Universiti Putra Malaysia (UPM)	187	Research	
4	Universiti Teknologi Malaysia (UTM)	184	Universities= 908 journals	66
5	Universiti Kebangsaan Malaysia (UKM)	158		
6	Universiti Teknologi Mara (UiTM)	87		
7	University of Multimedia (MMU)	81	Non-Research Universities=473 Journals	34
8	Universiti Teknologi PETRONAS (UTP)	78		
9	International Islamic Universiti Malaysia (IIUM)	77		
10	University of Nottingham Malaysia Campus (UNMC)	61		
11	MONASH Universiti Sunway Campus (MONASH)	51		
12	Universiti Tenaga Nasional (UNITEN)	38		
	Total	1381		100

The RU universities are more bound to published in IF journals to get more research funding. These universities receive a big amount of budget for R&D purposes and have to face pressure and make policies accordingly (<http://www.hir.um.edu.my>), and this is especially prevalent in Asian countries (Leydesdorff, 2009). The first five public universities (RU) published in 150-200 journals. Comparatively the private universities had fewer publications and published in 50 to 100 journals. The average number of journals for RU and non-RU universities is 182 and 68 respectively.

Analysis and Findings

Exploratory Factor Analysis (EFA)

In a tie with the problem, this section proceeds accordingly with descriptive statistics, data normality and EFA for our set of indices as presented in Table 3.

Descriptive Statistics and Normality Analysis of Complete Dataset

Descriptive statistics along with Skewness and Kurtosis are presented in Table 4. The results of the normality test based on raw data (excluding outliers) are reported in Table 5. The Skewness and Kurtosis are valid tests to find the normality of data. Their values show a normal distribution of data adequately normal. Keeping in view the requirement of EFA statistical application we used two other options as well. We also examined the relation between the raw, logarithmically transformed shifted ($\ln(x + 1)$) and square root transformation.

Table 5 shows a better Kaiser-Meyer-Olkin (KMO) results and a slight better-explained variance for log data. For this reason, we found the logarithmic transformed data more adequate for EFA. Bornmann, Mutz and Daniel (2008; 2009) used a cut-off threshold >0.6 for extraction loading factors while Schreiber, Malesios and Psarakis (2012) fixed it at > 0.685 for Varimax rotation.

Schreiber *et al.* (2012) argued that small sample size for EFA can produce reliable results. Quite a few factors and high communalities are in favour of small sample sizes (Preacher and MacCallum, 2002). Further, to measure a sampling adequacy, a specific test Kaiser-Meyer-Olkin (KMO) of value >0.5 is acceptable (Kaiser, 1974). KMO value (Table 6) of the present data sample is >0.5 with high communalities (>0.85) (Table 7). Based on KMO values and variance explained (Table 6 and 7), we finally utilized logarithmically transformed data. We identified two unknown factors through Eigen values (>1) via variance explained.

This is evident that EFA can be used and is appropriate for our formulated problem and dataset. Initially, we considered eleven indices, TP, TC, IHI and 8 of JRI (JIF, CIF, IF avg, MIF, CIF, CIF5Y, Imm-Index, and EF). This set of indices produced inadequate results for EFA. After omitting the TP, we applied EFA to TC, IHI, and 8 JRI (IF, CIF, IFavg, MIF, CIF, CIF5Y, Imm- Index, and EF).

Table 3. Analysis of Complete dataset for institutional level indices applied

University	TP	TC	IHI	JIF	CIF	AIF	MIF	IF(5Y)	CIF(5Y)	Imm-Index	EF
USM	724	4027	26	311.36	1609.71	2.229	1.35	331.43	1705.82	49.752	2.506
UPM	551	2309	20	255.12	879.04	1.600	1.12	262.86	886.18	40.100	2.070
UM	495	2388	23	337.45	948.07	1.950	1.50	318.54	871.69	52.598	2.481
UTM	475	2259	23	262.16	883.14	1.883	1.12	280.76	910.61	39.835	2.277
UKM	386	1490	17	233.65	624.13	1.634	1.25	246.65	629.14	36.081	1.975
UiTM	139	359	9	144.85	239.58	1.815	1.39	154.08	248.73	21.922	1.318
IIUM	138	251	7	100.01	174.87	1.270	1.02	103.96	177.20	14.640	0.960
MMU	532	2231	19	120.22	583.83	1.099	1.17	128.66	576.70	18.130	0.874
UNMCC	126	616	13	102.82	248.58	1.973	1.55	100.34	241.58	15.450	0.776
UTP	142	329	9	122.97	263.12	1.853	1.31	134.24	287.38	19.896	1.179
MONASH	76	302	10	87.87	131.94	1.713	1.59	94.86	140.93	13.533	0.887
UNITEN	71	139	6	50.86	91.77	1.293	1.22	55.65	100.24	7.460	0.351

Analysis of EFA

Table 6 reports the results of KMO values of the transformed data for the appropriateness of factor analysis. The next table 7 reveals the results of communalities for 3 EFA models that are the “variance in observed variables accounted for by a common factor” (Hatcher, 1994).

Table 4. Descriptive statistics

<i>Indices</i>	<i>Descriptive Statistics</i>					<i>Skewness</i>	<i>Kurtosis</i>
	<i>Mean</i>	<i>St.dev</i>	<i>Median</i>	<i>Min</i>	<i>Max</i>		
TP	321.25	229.079	264.00	71	724	0.364	-1.47
TC	1391.67	1246.835	1053.0	139	4027	0.776	-0.17
IHI	15.17	7.004	15.00	6.00	26.0	0.151	-1.60
IF	177.44	96.683	133.90	50.85	337.45	0.452	-1.34
CIF	556.48	457.445	423.47	91.77	1609.7	1.115	1.02
MIF	1.30	0.182	1.28	1.02	1.59	0.239	-1.01
AIF	1.69	0.332	1.76	1.10	2.23	-0.427	-0.44
IF(5Y)	184.34	97.047	144.15	55.65	351.43	0.351	-1.58
CIF(5Y)	564.68	471.04	432.04	100.24	1705.8	1.317	1.87
Imm-index	27.45	15.356	20.91	7.46	52.60	0.471	-1.32
EF	1.47	0.748	1.249	0.35	2.51	0.179	-1.56

Overview of Statistical Procedure for EFA

Table 5. Test for normality of data

	<i>Kolmogorov-Smirnov^a</i>			<i>Shapiro-Wilk</i>		
	<i>Statistic</i>	<i>Df</i>	<i>Sig.</i>	<i>Statistic</i>	<i>df</i>	<i>Sig.</i>
TP	.283	12	.009	.863	12	.053
TC	.233	12	.071	.852	12	.038
IHI	.186	12	.200*	.918	12	.267
IF	.208	12	.158	.881	12	.090
CIF	.235	12	.067	.856	12	.043
AIF	.183	12	.200*	.929	12	.369
MIF	.114	12	.200*	.960	12	.782
IF(5Y)	.228	12	.085	.876	12	.078
CIF(5Y)	.212	12	.143	.829	12	.020
Imm-index	.228	12	.086	.904	12	.178
Eigen Factor	.180	12	.200*	.937	12	.458

*At a 5% Significance Level

Table 6. Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy

	<i>X</i>	$\sqrt{x}$	$\ln(x + 1)$
KMO	0.564	0.540	0.695
Sig.	0.00	0.00	0.00

Table 8 provides Initial Eigenvalues >1 and indicates that the total variance explained by first two factors is 75%, and 17 % of cumulative variance explained by both factors are 91%. Component matrix (Table 8) illustrates that the set of indices clearly loads on two extracted factors. Rotated Component Matrix Table (9) for EFA model shows that the indices have

substantial loading on two established factors. It indicates the loading of two institutional ‘impact of the productive core indices’ (TC and IHI) and six others JRI have high loading (> 0.90) and a slight less for EF (>0.891).

Table 7. Communalities for 3 EFA models

	<i>X</i>		$\sqrt{x}$		$\ln(x + 1)$	
<i>Indices</i>	<i>Initial</i>	<i>Extraction</i>	<i>Initial</i>	<i>Extraction</i>	<i>Initial</i>	<i>Extraction</i>
TC	1	0.893	1	0.9	1	0.896
IHI	1	0.883	1	0.877	1	0.866
IF	1	0.94	1	0.951	1	0.953
CIF	1	0.934	1	0.958	1	0.962
IF(avg)	1	0.854	1	0.865	1	0.841
MIF	1	0.869	1	0.844	1	0.87
IF(5Y)	1	0.954	1	0.963	1	0.967
CIF(5Y)	1	0.879	1	0.925	1	0.950
Imm- Index	1	0.918	1	0.943	1	0.955
EF	1	0.869	1	0.861	1	0.870

AIF and MIF both have substantially high loading on the second factor>0.9. MIF is more accurate measure than the average value, due to the impact factor’s skewed distribution (Costas & Bordons, 2007). IF and CIF and IF5y and CIF5y require two years and five years time span with different strengths of productivity. EF is another index based on 5-year data excluding journal self-citation to rate the total importance of journal. Journals generating higher impact on the field have larger Eigenfactor scores (Bergstrom, 2007). “EF improves upon JIF and somewhat robust indicators of quality and prestige of the journal due the inclusion of 5 year's data, exclusion of journal self-citations” (YangYin, 2010, p.3). Rather a high journal EF depicts producing of high-impact scientific findings in a specific area (YangYin, 2010; Saad, 2006). IF (5y) indicates the speed with which citations to a specific journal appear in the published literature. Immediacy index that is based on one-year data shows the same value as CIF on the first factor. They both require a different strength of data. Surprisingly they all loaded on the same factor along with IHI.

Table 8: Total variance explained for 3 EFA models.

<i>Data type</i>		<i>Initial Eigenvalues</i>			<i>Extraction Sums of Squared Loadings</i>			<i>Rotation Sums of Squared Loadings</i>		
		<i>% of</i>		<i>Cumulative %</i>	<i>% of</i>		<i>Cumulative %</i>	<i>% of</i>		<i>Cumulative %</i>
		<i>Total</i>	<i>Variance</i>		<i>Total</i>	<i>Variance</i>		<i>Total</i>	<i>Variance</i>	
Raw indices	1	7.401	74.006	74.006	7.401	74.006	74.006	7.269	72.687	72.687
	2	1.594	15.940	89.946	1.594	15.940	89.946	1.726	17.259	89.946
$\sqrt{x}$	1	7.432	74.325	74.325	7.432	74.325	74.325	7.314	73.142	73.142
	2	1.655	16.547	90.872	1.655	16.547	90.872	1.773	17.730	90.872
$\ln(x+1)$	1	7.457	74.569	74.569	7.457	74.569	74.569	7.343	73.427	73.427
	2	1.672	16.720	91.290	1.672	16.720	91.290	1.786	17.862	91.290

Table 9. Rotated component matrix

<i>Indices</i>	<i>Components</i>	
	<i>1</i>	<i>2</i>
C	.945	-.055
IHI	.929	.059
IF	.965	.147
CIF	.978	-.074
AIF	-.133	.907
MIF	.309	.880
IF(5Y)	.970	.159
CIF(5Y)	.974	-.038
Imm-index	.950	.230
EF	.891	.275
Eigenvalues	7.401	1.595
Variance explained	75%	17%

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

Values >.5 are bold.

Conclusions

The caveats of h-index, JIF, and traditional metrics have been discussed in the abundant literature. Previous studies are meaningful to understand the properties of newly introduced indices and potential use of Institutional's h-index as a complement aid with IF and its variants. (Bador & Lafouge, 2010; Bornmann et al., 2012; Yang Yin, 2011) or, as a supplement (Braun, Glanzel & Schubert, 2006).

The present study describes the case of Malaysian engineering research applying the scientometric approach, method and techniques for RPE. Based on the ten years data analysis from WoSTM, we applied a set of comparatively new indices. To achieve the research objectives, empirical analyses were carried out, and hypotheses were examined statistically.

The major findings of the study demonstrate that there seems to be increasing the trend to get published in IF journals. A steady increase of IF publications is observed from 2001 in the Malaysians scientific productivity of all studied disciplines including engineering. The ambition to publish in IF WoSTM recognized publications is reinforced by the Malaysian Research Assessment (MyRA) exercise, which requires institutions to publish papers that are indexed in the citation database. This is due to the Malaysian Ministry of Education policies towards research and publications during two five years plans (2001-2005; 2006-2010). RU status universities (shared 68% and 74% publications and citations). These universities have published in 66% of total journals. Overall, the RU universities lead in positioning order with the application of indices. USM is an exceptional case and remained in position one with respect to almost all indicators. While others showed a noteworthy change in their positioning order. IHI has stronger functional relation with institutional citation data followed by publication record. Institutional citation data is the best predictor of IHI. Often used metric C (as total impact indicator) and the EF (as prestige indicator) have a high association with IHI. This establishes the property of h- index as prestige impact measure of scientific productivity. This index appears a useful yardstick, because of good functional relationship with C and P and has shown some discriminatory power for ranking purpose. The EFA suggests the same distinguishing behaviour of IHI like P and C. The findings put forward a better understanding

about the consideration of new impact metric for RPE at the meso level. Malaysian engineering institutional case indicates that h-index and others metric have not only strong association for total institutional citation data but also with institutional cumulative journal indices. However, the total variance explained for two components yields about 75% for its first component and 16% for the second component. Therefore, findings are based within the limitations of the statistical analysis.

Publishing in high-quality IF journals is important if a country is to realize its ambition to have its universities amongst the top rated universities in the world. This is not peculiar to Malaysia. The Ministry of Education Malaysia is targeting two research universities in the country to be in the top world 100 best universities by 2020. Other countries also place a high emphasis on publishing in IF journals and would want to be ranked as top world universities, even if they are not always explicit in saying so. Given the significant number of papers that have now been published by Malaysian institutions (56, 571 in Web of Science, Essential Science Indicators, Web of Science 2015), there is an opportunity to carry out further analysis. It would be interesting, for example, to provide analysis at a discipline level to get a feeling for the strengths of the institution at a lower level. It would also be informative to consider other normalization measures to ascertain if they provide a better correlation with the MyRA ranking.

Acknowledgement

The work of Muzammil Tahira and A. Abrizah was supported by the Ministry of Higher Education Malaysia (HIR-MOHE) UM.C/HIR/MOHE/FCSIT/11.

References

- Amin, M. & Mabe, M. (2000). Impact factor: Use and abuse. *Perspectives in Publishing*, 1, 1-6.
- Bador, P. & Lafouge, T. (2010). Comparative analysis between impact factor and h-index for pharmacology and psychiatry journals. *Scientometrics*, 84(1), 65-67.
- Bornmann, L. Marx, W., Gasparyan, A.Y., & Kitas, G. D. (2012). Diversity value and limitations of the journal impact factor and alternative metrics. *Rheumatol International*, 32, 1861-1867.
- Bornmann, L., Mutz, R., Hug, S. E., & Daniel, H.D. (2011). A multilevel meta-analysis of studies reporting correlations between the h index and 37 different h index variants. *Journal of Informetrics*, 5(3), 346-359.
- Bornmann, L., Mutz, R., & Daniel, H.D. (2009). Do we need the h-index and its variants in addition to standard bibliometric measures? *Journal of the American Society for Information Science and Technology*, 60(6), 1286-1289.
- Bornmann, L., Mutz, R., & Daniel, H.D. (2008). Are there better indices for evaluation purposes than the h-index? A Comparison of nine different variants of the h-index using data from biomedicine. *Journal of the American Society for Information Science and Technology*, 59(5), 830-837.
- Braun, T., Glanzel, W., & Schubert, A. (2006). A Hirsch-Type index for journals. *Scientometrics*, 69(1), 169-173.
- Conway, J.M. & Huffcutt, A. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods*, 6(2), 147-168.
- Leydesdorff, L. (2009). How are new citation-based journal indicators adding to the bibliometric toolbox? *Journal of the American Society for Information Science and Technology*, 59(2), 278-287.
- Mingers, J., Macri, F., & Petrovici, D. (2012). Using the h-index to measure the quality of journals in the field of business and management. *Information Processing and Management*, 48,234-241.
- Moussa, S. & Touzani, M. (2010). Ranking marketing journals using the Google Scholar-based hg-index. *Journal of Informetrics*, 4, 107-117.
- Prathap, G. (2011). Correlation between h-Index, Eigenfactor™ and Article Influence™ of chemical engineering journals (Letter). *Current Science*, 100(9), 1276.
- Raj, R.G. & Zainab, A.N. (2012). Relative measure index: A metric to measure quality. *Scientometrics*, 93(2), 305-317.
- Schreiber. Malesios, C. C. & Psarakis, S (2012). Exploratory factor analysis for the Hirsch Index, 17 h-type variants, and some traditional bibliometric indicators. *Journal of Informetrics*, 6, 347- 358.

- Schreiber, M., Malesios, C. C. & Psarakis, S. (2011). Categorizing Hirsch Index variants. *Research Evaluation*, 20(5), 397–409.
- Thomson-Reuters. (2015). Web of Science™ Retrieved on March 20, 2015 from http://admin-apps.webofknowledge.com/JCR/help/h_impfact.htm
- Yang Yin, C. (2011). Do impact factor, h-Index and Eigenfactor™ of chemical engineering journals correlate well with each other and indicate the journals' influence and prestige? *Current Science*, 100 (5), 648-653.

Factors Influencing Research Collaboration in LIS Schools in South Africa

Jan Resenga Maluleka¹, Omwoyo Bosire Onyancha², Isola Ajiferuke³

¹ *malulrj@unisa.ac.za* ² *onyanob@unisa.ac.za*

University of South Africa, Dept of Information Science, PO Box 329 Unisa, 0003

³ *iajiferu@uwo.ca*

University of Western Ontario London, Ontario, Canada, N6A 5B7

Abstract

The study sought to explore the underlying factors that influence research collaboration in Library and Information Science (LIS) schools in South Africa. The population for the study consisted of 85 academic teaching staff employed by LIS schools in South African universities. A survey design was used to obtain data for the study, through a questionnaire containing open- and close-ended questions. A total of 85 teaching staff in 10 LIS schools in South Africa were alerted, through email, to the location of the Web-based questionnaires, developed using the Stellarsurvey software. A total of 51 questionnaires were completed and returned for analysis. The findings suggest that factors such as networking, sharing of resources, enhancing productivity, educating students, overcoming intellectual isolation, and accomplishments of projects in a short time as well as learning from peers influenced research collaboration in LIS in South Africa. Factors that are likely to hinder effective collaboration in LIS research include bureaucracy, lack of funding, lack of time, as well as physical distance between researchers. The findings further suggest that even though there are drawbacks to collaboration, majority of LIS researchers thought that collaboration is beneficial and should be encouraged.

Conference Topic

County-level studies

Introduction

In today's global economy, there is an increasing importance of collaborative relationships between individuals, organisations, and even countries. Collaboration, defined as a "process where two or more individuals or organizations deal collectively with issues that they cannot solve individually" (Ocholla, 2008:468) and "the working together of researchers to achieve the common goal of producing new scientific knowledge" (Katz & Martin, 1997), can be found in all the spheres of human life, for example in politics, economics or even in religion. Katz & Martin (1997) are of the opinion that research collaboration has significant benefits such as intellectual championship, joint development of skills, effective transfer of knowledge and the improvement of potential visibility of researchers. For example, collaboration can build partnerships and help empower researchers to accomplish projects that were never going to be easy to do individually. Collaboration brings together experiences, skills, knowledge and the know-how of different researchers into one particular project. By way of research collaboration, researchers from different countries (both developed and developing countries) come together for different purposes, among which are sharing of information, knowledge and technological transfer as well as finding solutions to specific problems (Onyancha, 2009). Researchers collaborate in order to accomplish tasks that cannot be accomplished as isolated individuals. Onyancha & Ocholla (2007), too, note that securing research grants is to a large extent becoming increasingly pegged on whether the intended research would be conducted collaboratively. Collaboration can be important especially in developing countries where there might be a lack of scientists and resources in certain fields. The few available researchers in developing countries can collaborate with those in developed countries for the former to be active in research as well as flourish as scientists.

According to Katz and Martin (1997), collaboration among scholars in both natural and social sciences has been steadily increasing for decades, covering different disciplines, development categories, institutions, geographic regions and countries. The increasing attention on research collaboration in LIS has also been pointed out by Onyancha and Maluleka (2011). Sugimoto (2011) argues that research in the field of LIS has followed similar patterns of increased collaboration as in other fields. According to Ocholla (2008), collaboration and partnerships could be forged amongst LIS institutions in a country and internationally or regionally in areas such as teaching, research, student and staff exchange, conferences and workshops, curriculum development, publications, research supervision and examination and distance teaching/research.

Rationale for the study

An examination of the published literature reveals that several studies have been conducted to examine research collaboration in different fields or disciplines including LIS. The focus of these studies includes identifying the collaborating authors, institutions, and/or countries (e.g. Sun, 2006; Onyancha & Ocholla, 2007), measuring the strengths of research collaboration (e.g. Yamashita & Okubu, 2006) and examining the nature of collaboration (e.g. Katz & Martin, 1997; Smith & Katz 2000). Several other studies have majorly focused on answering the question ‘who’ or ‘what’ of collaboration. In other words, studies that have been conducted previously on collaborative research have largely focused on the frequency of collaboration between the authors, the nature of collaboration and the strength of collaboration across disciplines. To the best of the researchers’ knowledge, little has been done to answer the question ‘why?’ The current study therefore aims to investigate those factors that may influence collaboration in LIS schools in South Africa. The main objective of this study is to find out the underlying reasons and/or factors that influence collaboration, a situation that may explain the quantitative results (e.g. trends, patterns, and type of research collaboration) reported in previously published works.

Research Questions

The following research questions were posed in order to fulfil the study’s main objective;

- What factors hinder and/or would hinder effective research collaboration in LIS schools in South Africa?
- What factors do and/or are likely to foster effective research collaboration in South African LIS schools?
- To what extent do the enhancers and inhibitors of collaboration influence research collaboration in LIS schools in South Africa?

Methodology and Materials

The study adopted a survey design to seek for the LIS academics’ views on factors that influence research collaboration in LIS research in South Africa. Neuman (2007:273) argues that survey research is developed within the positivist approach and it is the mostly and widely used design in the social sciences. Similarly, Leedy and Ormrod (2010:187) argue that survey research *involves acquiring information about one or more groups of people – perhaps about their characteristics, opinions, attitudes, or previous experiences by asking them questions and tabulating their answers.*

In this study, the survey involved all academic teaching staff employed by LIS schools in South African universities. They include teaching assistants, junior lecturers, lecturers, senior lecturers, associate professors, and professors. Honorary professors, research fellows, extraordinary professors, or any other scholars who are linked to a particular department but without being fulltime were excluded as they appeared to have more than one institutional

affiliation. With only ten LIS schools offering LIS education in South Africa, there was no sampling conducted as all schools were included in the study. The total number of the teaching staff was also small, leading us to include all academics in the target population for this study. Table 1 shows the number of staff in the LIS departments by the parent University.

Table 1. LIS Schools in South Africa

School name	Acronym	Number of teaching staff
University of South Africa	UNISA	19
University of Pretoria	UP	24
University of KwaZulu-Natal	UKZN	6
University of Zululand	UZ	7
University of Fort Hare	UFH	4
University of Cape Town	UCT	8
University of the Western Cape	UWC	6
Durban University of Technology	DUT	5
University of Limpopo	UL	4
Walter Sisulu University	WSU	2
TOTAL		85¹

The instrument of data collection for the study was a questionnaire, which was deemed to be the most appropriate. The questionnaire contained both closed-ended and open-ended questions, the former being the majority. There were a total of 20 questions focusing on specific items that were linked to the research questions. We used the “*Stellarsurvey*” online survey software as a platform for the questionnaires.² We then sent emails to all the identified LIS researchers in South African LIS schools. The emails contained a link directing them to the website which invited them to participate in the study. Respondents were given three weeks to complete the questionnaire online. After three weeks a reminder was sent to participants again reminding those who had not responded to do so.

Results and discussion

Profile of the respondents

Out of the 85 teaching staff members that were approached to participate in the study, only 51 completed the questionnaires, leading to a response rate of 64.6%. It was found that 43% (i.e. 22) of the respondents were male while 29 (57%) were female. All respondents had a university qualification ranging from a bachelor’s degree to doctoral degree. The majority of the respondents (i.e. 21 or 41%) had a master’s degree as their highest qualification, followed by those with a doctoral degree (i.e. 19 or 37%) and then those with honours (11 or 22%). The majority of the respondents were employed as lecturers (27 or 54%), followed by junior lecturers (9 or 18%) and full professors (5 or 10%) while senior lecturers and associate professors stood at 3 (3%) each. The results shows that the majority of the respondents are actively involved in research either as masters and doctoral students or as supervisors and mentors for these students.

¹ The number of the teaching staff was retrieved from the LIS departments’ websites.

² The software is available at: <http://stellarsurvey.com/>.

The status of collaboration in LIS research

It was found that 43 (84%) of the respondents collaborated in the conduct of research while only 8 (16%) indicated that they never collaborated before. The results in Figure 1 (a) reveal that 45 (88%) respondents believe and agree that collaboration in research is important while 2 (4%) were neutral with only 4 (8%) saying collaboration in research is not important.

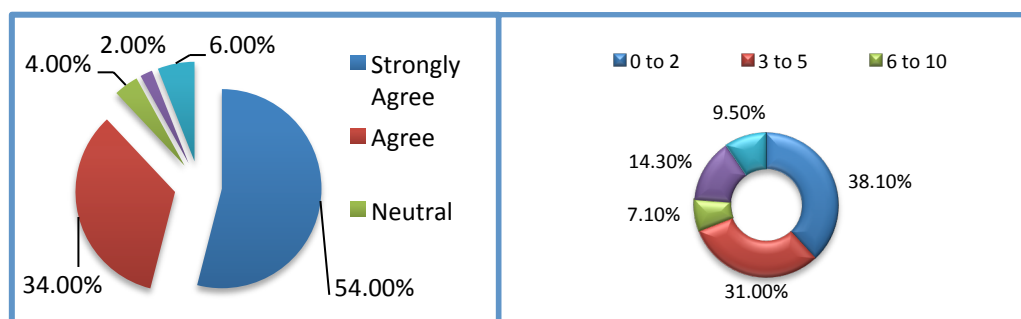


Figure 1. (a) Importance of collaboration (N=51) (b) The number of collaborated projects that are already published.

It is strange to note that while 84% of the respondents indicated that they collaborated, there was a sizable number, who may have included the ones who reported that they collaborated, who might have felt that collaboration is not important. This group could include researchers who are forced, by circumstances (e.g. institutional policies on co-supervision of students or mentorship of junior colleagues). When we looked at collaborative projects already completed (Figure 2 (b)), 32 (62%) respondents had already completed three or more projects collaboratively while only 19 (38%) had completed between 1 and 2 projects collaboratively. It was worth noting that the current generation of researchers are actively engaged in collaborative research. Results tend to imply that the researchers prefer sharing and working together as compared to the past where the degree of collaboration among researchers has been reported to be low.

It has been shown that research collaboration in South Africa has increased tremendously in the previous decade (i.e. 2001-2009) (Sooryamoorthy, 2009). There are a number of reasons that may have influenced this pattern on collaborative research. Universities in South Africa have realised that they are losing their most experienced researchers who were approaching retirement age before the young developing researchers were fully equipped in the area of research. In some universities such as UNISA, huge funds have been invested into the development of young researchers through initiatives such as the mentorship programmes. This is done in view of Liebowitz's (2009) suggestion that formal mentoring programmes are popular techniques used for knowledge sharing, knowledge retention, knowledge transfer, and also to enhance worker skills. In this programmes, senior researchers are assigned mentees who learn from them on a daily basis for a specific period of time. Research funding organisations such as the National Research Foundation (NRF) of South Africa are also making funds available for collaborative and multidisciplinary research. Doctoral students are also funded to conduct post-doctoral research in collaboration with their mentors. The responses from the questionnaire also suggest that other universities have made it compulsory for supervisors to publish at least one article collaboratively with their students from the latter's theses and dissertations. The above is evident from the feedback from the respondents and it may be the reason why the majority of the respondents in the survey indicated that they are engaged in collaborative research, although some of them also indicated that collaboration is not important.

Looking at the group of people that the respondents mostly collaborated, it was noted that the researchers in LIS schools in South Africa largely collaborate with fellow researchers when taking the occasional, often and most often times of collaboration into account; the three account for 80% (see Table 2). This suggests that LIS researchers prefer collaborating with fellow researchers, preferably in their own field of interest. The main reason may be that working on a project with someone who understands one's subject area and the methodologies involved may result in the project being completed at a faster pace than if the opposite had to happen.

Another point worth highlighting is the results on collaboration with international researchers which was very low, with over 70% of the respondents indicating that they never collaborated at this level. This pattern is contrary to previous studies' findings, which revealed that most research in Africa is published in collaboration with international researchers (see Narvaez-Berthelemot, Russell, Arvanitis, Waast, & Gaillard, 2001). It is therefore unfortunate to find that researchers in LIS schools largely collaborate locally as opposed to engaging in international collaboration as researchers collaborating at the international arena have a competitive advantage over their peers because they have a chance of using resources from both institutions to which they are affiliated. The other notable advantage worth mentioning about international collaboration is the fact that it allows researchers a chance to publish in international journals, share international experiences which will allow them an opportunity to gain international visibility. Narvaez-Berthelemot, Russell, Arvanitis, Waast, & Gaillard (2001) note that researchers in developing countries would also benefit from their peers in developed countries in terms of publication of their research in international journals. The authors opine that *"the less productive the developing country, the greater the dependence on international co-authorship for mainstream publication"*. Katz and Martin (1997) observe that most governments have been keen to increase the level of international collaboration engaged in by the researchers whom they support in the belief that this will bring about cost-saving or other benefits. The main reason given by respondents for not collaborating at this level was distance and logistical problems that exist when working with someone from another country. The other reason worth noting is the fact that researchers from bigger institutions or developed countries may undermine the contribution of the other researchers from poorer countries or smaller institutions. The opposite may also happen where researchers from smaller institutions may lack self-belief, contribute less and end up not playing an equal role in the whole collaborative venture.

Table 2. Group of persons that respondents collaborated with

	Never	Rarely	Occasionally	Often	Most often
Students	33.3%	7.7%	25.6%	23.1%	10.3%
Mentor	24.3%	18.9%	13.5%	16.2%	27.0%
Mentees (other than students)	50.0%	14.7%	20.6%	11.8%	2.9%
Fellow Researchers	5.0%	15.0%	30.0%	45.0%	5.0%
Senior Researchers	28.2%	15.4%	15.4%	20.5%	20.5%
International Researchers	45.9%	24.3%	10.8%	13.5%	5.4%

It seems like there is need for institutions to initiate programmes geared towards supporting the researchers in overcoming problems faced during international collaboration. The researchers also need to take advantage of the latest technologies that can easily allow them to work together without having to travel between countries. For LIS researchers in South Africa to remain at par with their international counterparts, they need to engage with them and work with them collaboratively so that they don't work in isolation.

Table 3. Groups likely to collaborate with in the future

	Never	Rarely	Occasionally	Often	Most often
Students	2.6%	7.9%	23.7%	39.5%	26.3%
Mentor	24.3%	16.2%	16.2%	21.6%	21.6%
Mentees(other than students)	25.7%	14.3%	34.3%	20.0%	5.7%
Fellow Researchers	0.0%	12.2%	22.0%	43.9%	22.0%
Senior Researchers	12.5%	20.0%	10.0%	35.0%	22.5%
International Researchers	12.5%	20.0%	30.0%	20.0%	17.5%

Enhancers and Impact of collaboration

Merlin (2000), Katz and Martin (1997), Bozeman and Corley (2004) give a summary of the following factors that are likely to foster effective collaboration in research:

- Collaborative research allows young researchers, access to expertise /experts with specialised knowledge and expertise in a particular area and learns directly from them.
- These partnerships gives researchers an opportunity to share resources where researchers from smaller institutions will get access to resources from big institutions and again institutions to supplement each other
- Multidisciplinary research allows a cross pollination of ideas and collaborative research allows partners to learn from one another
- There are more chances of getting funds if a collaborative initiative is submitted to funding organisation. Secondly a project can get funds from both organisations with will make it possible to carry out
- Working alone in a particular project can make one feel lonely and isolated. Working in a team helps one to overcome that intellectual isolation.

For this study, respondents were asked to indicate the extent to which factors such as networking, sharing of resources, enhancing productivity, educating students, overcoming intellectual isolation, accomplishment of projects in a short time, learning from peers, and incentives influence them (researchers) to engage in collaborative research.

The results indicated that over 44 (86%) respondents engage in collaborative research to strengthen their networks with other scholars. The respondents reported that networking helps to bring these scholars who happen to have common interests together and create partnerships that often last for longer. Researchers usually work alone on their projects which leaves them isolated. Networking or coming together with fellow researchers to work on a project together may help overcome that isolation. The importance of networking was also highlighted by 37 (73%) respondents who indicated that they collaborate in research to overcome intellectual isolation. Another patch of respondents numbering 38 (75%) also agreed to be collaborating with an aim of sharing resources. This can be very significant to researchers from smaller institutions and underdeveloped countries with little resources. Such partnerships can allow them to take advantage of the available resources in both institutions, some of which may not be available in their smaller institutions.

Learning from peers was also one of the most common factors among respondents on why they collaborate in research. The results show that 43 (84 %) respondents collaborate in research to learn from their peers. This usually happens where two or more scholars with different expertise come together to solve a research problem. Each researcher brings a special skill that may not be known by the others and that brings an opportunity for all to learn from one another. There were mixed feelings among respondents when it came to having to collaborate to get incentives. In South Africa, a number of institutions usually attach incentives to publications published in selected peer reviewed journals, book chapters, peer reviewed conference proceedings and books that earn subsidy from the Department of Higher

Education and Technology (DoHET). Only 24 (47%) respondents indicated that incentives may influence them to collaborate with 21 (41%) saying incentives have very little influence on them when it comes to collaborating. It has been informally noted by researchers at some forums of discussion that some researchers at times choose not to collaborate so that they don't share incentives made available and opt to work alone. This can have serious implications because those who are skilled enough will work alone and continue getting incentives while they are not leaving anyone to take over from them when they retire which will create a knowledge gap. Having incentives for research in an academic setting is motivating and encouraging for researchers but it has negative implications for the future.

Reasons for collaborating

Respondents were requested to give specific reasons that are likely to foster collaborative initiatives with particular groups such as, students; mentors; mentees (other than students); colleagues in the same department; fellow researchers; and international researchers.

Reasons for collaborating with students and mentees (other than students)

The responses received for this question were not that surprising considering the population for this study. Respondents indicated that they collaborate with students to impart knowledge and help the latter to obtain their qualifications. Some respondents indicated that collaborating with students is part of their jobs. A number of promoters feel that it takes a lot of time to do postgraduate supervision and as a result, they make sure that they get an article out of the whole project so that their efforts do not go to waste. It was also interesting and encouraging to note that some supervisors feel that students bring fresh perspectives on themes and ideas that they may be having at the time. This means that such supervisors give students a platform and opportunity to participate in the whole project while taking their ideas into consideration. Furthermore, respondents indicated that they would like to share their experiences on a particular subject and help capacitate their mentees while strengthening their relationships with their students at the same time exploring areas outside their subject specialisation.

Reasons for collaborating with mentors and managers

There was a general consensus among those respondents, who are being mentored by senior colleagues, that it is important to tap into the mentor's experience and knowledge in order to develop skills and research avenues. Mentorship of young researchers where the latter learns from the senior and experienced colleagues is again at the centre stage. Field (2001:270) is of the opinion that a mentor should play an important role in the career development of mentees, by providing them with background information and support for individual growth, as well as making them aware of opportunities available.

The other important thing about having a mentor is the creation of an opportunity to connect with the mentor's professional networks. This allows the mentee to grow and expand his/her professional boundaries. Mentorship can either be formal or informal. The best example of a formal mentorship is that of a supervisor working with a post graduate student. Informal mentoring may happen between the experienced and the less experienced through a personal connection. One respondent mentioned that mentors know their mentees best, and it is advantageous to work with someone who knows and understands his/her mentee well. Having worked with someone before gives the mentee an advantage of knowing how the mentor does things and what the latter expects of him/her. This is important during collaboration where responsibilities are shared because it will be helpful in deciding which role should be played by whom. Other respondents indicated that a natural consequence of being a young researcher and wanting to learn definitely motivated them in the conduct of collaborative research with their mentors.

Reasons for collaborating with colleagues in the same department

Being in the same department will most likely mean that one knows and understands each other's strengths and weaknesses. Respondents indicated that they collaborate with colleagues with the aim of producing high quality papers in a short space of time to enhance their productivity. Some respondents mentioned a desire to pursue niche areas in their departments as a reason for collaborating with fellow researchers. They indicated that such collaborative research has the potential to generate income for them and increase their research output. Some respondents indicated that they work on departmental joint projects and they have no choice or can't avoid them as they are in the same department. This group may not yield desired results because collaboration is not conducted between willing partners who are committed to seeing the project through to the end.

Other respondents mentioned that co-supervision of students' work automatically gets them to work together and eventually they publish together with the students. In view of the fact that some LIS schools in South Africa have closed down or changed focus to non-LIS disciplines, the onus is left to the few available LIS schools to ensure the survival of the profession. The closing down of LIS schools has put too much pressure on the few academics left in LIS as they are expected to service the increasing student numbers and also conduct research so they stay relevant. This situation encourages collaboration where researchers will share responsibilities and reduce the time and effort required to complete a task.

Reasons for collaborating with colleagues from other departments

The respondents indicated that collaborating with someone from another department in the conduct of research widens their horizons. The respondents further mentioned that such collaboration is very important because it helps with the establishment of interdisciplinary networks and exposure to a wide variety of research methods. The other notable reason mentioned by the respondents is the cross-pollination of ideas that will result from collaborating with someone from a different department or discipline.

Reasons for collaboration with International Researchers

This type of collaboration as discussed in the sections above enables researchers to share international experiences, foster international networks, and can help researchers do comparative studies with peers from other countries. Respondents who indicated that they have collaborated at the international level believe that global perspective is key to providing comprehensive research studies. Researchers can never work in isolation and the same should happen in LIS. International collaboration according to some respondents can increase researchers' chances of accessing funds and publications as well as get international visibility.

Barriers to collaboration

This section explores the issues that LIS scholars perceive to hinder effective research collaboration in LIS schools in South Africa. Katz and Martin (1997) gave a summary of the following barriers to collaboration:

- Financial implications in the form of travel costs , moving of equipment's and so forth
- Increased administration resulting from more people/institutions involved,
- Lack of time from some collaborators, or additional time required as different parts of the research will be done in different locations
- Different management cultures, financial systems and rules on intellectual property rights

Table 4. Barriers to collaboration

	To a great extent	Somewhat	Very little	Not at all
Bureaucracy	42.2%	33.3%	22.2%	2.2%
Lack of funding	43.5%	28.3%	19.6%	8.7%
Intellectual property rights	9.1%	29.5%	36.4%	25.0%
Lack of time	43.5%	28.3%	15.2%	13.0%
Clash of values	9.1%	31.8%	34.1%	25.0%
Ethics	15.9%	18.2%	27.3%	38.6%
Distance between researchers	15.2%	19.6%	23.9%	41.3%

For this study, respondents were first asked to indicate the extent to which barriers such as bureaucracy, lack of funding, intellectual property rights, lack of time, clash of values, ethics, and distance between researchers may have prevented them or are likely to prevent them from engaging in collaborative research. Secondly respondents were requested to indicate the extent to which a number of personal traits and characteristics may be a barrier/s to research collaboration. Table 4 provides the extent to which some factors act as barriers to effective collaboration.

Table 5. Personal traits or characteristics that may be a barrier to research collaboration

	<i>To a great extent</i>	<i>Somewhat</i>	<i>Very little</i>	<i>Not at all</i>
Gender	6.7%	15.6%	20.0%	57.8%
Level of education	31.1%	44.4%	20.0%	4.40%
Competencies	70.5%	29.5%	0.0%	0.0%
Honesty	72.7%	13.6%	6.8%	6.8%
Respect	80.0%	11.1%	6.7%	2.2%
Self-discipline	72.1%	23.3%	4.7%	0.0%
Work Ethic	75%	20.50%	4.5%	0.0%
Mutual Intent	75%	20.50%	4.5%	0.0%
Attitude	70.5%	25.0%	4.5%	0.0%
Interpersonal skills	47.7%	45.5%	2.3%	4.5%
Reliability	74.4%	23.3%	0.0%	2.3%
Nationality	4.7%	2.3%	20.9%	72.1%

A good majority of respondents (i.e. 39 or 76%) indicated that bureaucracy may be a barrier to collaboration. We believe that academics work under tight deadlines and the pressure to deliver is high and therefore too much red tape may sometimes delay their progress. Again over 36 (71%) respondents indicated that lack of funding maybe a barrier to collaboration. It should be noted that many institutions make funds available for research but if access to those funds is a problem then little research will be done. If a project does not receive funds then it will never get off the ground. It was interesting and surprising to note that 34 (66%) respondents indicated that ethics has very little impact on whether they collaborate or not. We opine that ethics is very important in research and perhaps that is why institutions around the world have adopted specific ethical principles when it comes to research. Only 17 (34%) respondents indicated that ethics may be a great barrier and influence their decision to collaborate. The distance between researchers also seem not to be a problem among respondents with 33 (65%) respondents indicating that it will not stop them from collaborating. The latest computer technologies such as Skype make it possible to work with someone who is in another country as if one were in the same room, so the issue of distance is increasingly becoming a thing of the past.

The majority of the respondents (i.e. 29 or 57.8%) did not see gender as barrier to collaboration. However someone's level of education was considered very important by the respondents. Over 38 (75%) respondents indicated that someone's level of education may be a barrier to collaboration. This may be influenced by the fact that researchers collaborate to accomplish goals that they cannot accomplish on their own; as a result, someone who is not academically capable may not be a good partner to have especially when one is under pressure to deliver. This was supported by the fact that all respondents suggested that somebody's inadequate competencies is definitely a barrier to collaboration. Personal characteristics such as honesty, respect, self-discipline, as well as attitude had over 46 (90%) respondents strongly indicating that the attributes will definitely block them from collaborating. Everybody wants to be associated with a well-mannered and respected person as well as someone who is not troublesome.

Reasons for not collaborating

Just like in the study by Katz and Martin (1997), this study investigated those underlying reasons that may hinder collaboration in LIS in South Africa. Respondents were asked to provide reasons that best describe why they may not collaborate with the following groups: students, mentors, Mentees other than students, colleagues in the same department, fellow researchers, seniors or managers and international researchers. The following were results as obtained from the survey.

Reasons for not collaborating with students and mentees

There was a general feeling amongst respondents that they will never work with students who are lazy and not prepared to work. This factor cannot be overemphasized as respondents mentioned issues like, lack of competencies, poor work ethic, and not following instructions on the students' side as main reasons they may not collaborate with students. Students who are repeating the same mistakes or not considering any advice or guidance given to them may be left without mentors. The respondents feel that such students may delay them at times as they do not stick to deadlines and agreements. Senior researchers may want to share their knowledge and skills but if the partner is not willing to learn then it defeats the whole purpose. Senior researchers are rated and evaluated according to their output and therefore wasting time on someone who does not want to learn or not willing to learn may be costly for them. Other responses included lack of mutual understanding, lack of commitment, time constraints as well as if the two parties do not share common research goals.

Reasons for not collaborating with mentors and managers

There were no surprises when it came to reasons why researchers will not collaborate with their seniors or managers in the conduct of research. A number of respondents were concerned about the fact that their mentors or seniors make them do all the work but equally share the credit which is somehow discouraging to them. Even though this is obviously unethical, it is common knowledge that some mentors abuse their positions and take advantage of their mentees. Young researchers will be expected to do all the work with little contribution from their more senior collaborating partner. Respondents further mentioned that mentors always demonstrate authority, lack empathy and never listen to their suggestions. Ignoring the contribution made by the more junior researchers may be demoralising and may result in the young researchers losing interest in conducting research because of the lack of self believe. Managers or mentors have an obligation to build as any form of advice or feedback is supposed to build as opposed to being too harsh. Many masters and doctoral students never complete their studies as some mentors give poor feedback or criticism that is aimed at breaking the students. Some of the respondents mentioned a lack of work ethic, lack of time, and not getting valuable advice or input from their mentors as other reasons for not collaborating with their mentors. Mentors normally have a lot of commitments, and a

collaborative project with a student may not be a priority to them, while the student's development and growth will be depending on it. This can therefore discourage students from wanting to collaborate with mentors.

Reasons for not collaborating with colleagues in the same department

This was a very interesting question and some of the responses given were somehow unexpected. Respondents mentioned that some colleagues have drawn their own conclusions about others which affect or influence their decision to collaborate. This is again a question of underestimating others and having one's own biased perceptions of others before they get to know them. That is a personal problem and has to do with everybody's personality and can only be solved over time, even though it poses challenges. Other respondents indicated that they will never collaborate with colleagues in their department because some colleagues never give their ideas a chance. This is a problem everywhere; colleagues who are mostly quiet may keep their ideas to themselves in such partnerships. Others are not good in expressing themselves and will mostly keep to themselves. This may result in ideas that end up being used although they are not the best, just because they came from the most vocal participants. One respondent indicated that in some instances, the most vocal colleagues may have a good command of the English language, while their ideas lack substance. Some of the other reasons raised include selfish colleagues, clash of ideas, competencies, attitude; lack of work ethic, and professional jealousy which was really unexpected. Some colleagues may feel that involving others in projects and working together may improve their profile and maybe become a threat to them in the work environment. Such colleagues end up being selfish and holding on to information and blocking their fellow colleagues. Others indicated they are so busy to an extent that they do not have time to do any other extra work, including collaborative research. Issues relating to office politics and intellectual property rights were also highlighted as possible reasons why some respondents do not enter into collaborative initiatives with fellow colleagues in the same department.

Reasons for not collaborating with fellow researchers

This question aimed to get responses on why LIS researchers are not collaborating or may not collaborate with fellow researchers in other departments as well as those in other universities. Many responses given were similar to the ones given in the immediate question above. However the issue of different research interests came out ahead of others. Even though many universities encourage multi-disciplinary research, researchers seem to prefer working with scholars who understand their area of interest and methodologies involved in the research, just to name but a few. Other reasons included unethical behaviour, time and distance between researchers, and different agendas among collaborating researchers.

Reasons for not collaborating with international researchers

Most of the barriers already indicated in the preceding questions were also mentioned here. Other reasons which were given by respondents regarding this question and are worth mentioning include distance and logistical problems, lack of communication, and topical issues, just to list a few. There is a general feeling from many local researchers that it is really not easy to work with someone who is very far especially in another country, even though the technologies available today make this possible and better than before.

Conclusions

The study by Sooryamoorthy (2009) revealed that collaboration in research in South Africa has been growing steadily over the years. This implies that, even though there are difficulties and drawbacks associated with collaboration in research, LIS researchers are mainly focusing in all the benefits that come with such partnerships and therefore engaging in collaborative research. It is important to mention that, even though the benefits of collaboration are evident,

the drawbacks cannot be ignored. A re-look at the enhancers and inhibitors of research collaboration suggests that the distance between researchers, past relationships and the institution of affiliation most influenced who collaborated with whom. The results imply that LIS researchers prefer partnering with colleagues who are nearer, mainly from the same institution. The collaboration networks suggest that issues discussed above have had a major impact on the current status of collaboration in LIS research in South Africa.

Collaboration links between supervisors and students are very much evident and seem to be the most influencing factor on research collaboration among LIS researchers in South Africa. It is also very encouraging to see some partnerships between senior researchers from different schools which is crucial for the growth and development of research in the field. Ocholla (2008) has observed that collaboration of LIS schools is weak and largely informal. This was very evident in the current study, too. Collaboration mainly happened between individuals while departments rarely collaborate hence there is no evidence of students from a particular university collaborating with their peers from other universities. This finding concurs with the views of Ocholla & Bothma (2007) who indicated that collaboration among LIS schools and researchers in such areas as "teaching, research, student and staff exchange, conferences, workshops, curriculum development, publications, research supervision, examination is very important yet very minimal".

Acknowledgments

This paper reports part of the findings of Mr Jan Malulaka's Masters dissertation (University of South Africa, 2014), supervised by Prof Omwoyo Bosire Onyancha, Chair and Professor, Department of Information Science and Prof Isola Ajiferuke, University of Western Ontario.

References

- Bozeman, B. & Corley, E. (2004). Scientists collaboration strategies: Implications for scientific and technical human capital. *Research Policy*, 33(4), 599-616
- Field, J. (2001). Mentoring: a natural act for information professionals? *New Library World*, 102(7/8), 269-273.
- Katz, J.S. & Martin, B.R. (1997). What is research collaboration? Retrieved February 20, 2010 from: http://www.sussex.ac.uk/Users/sylvank/pubs/Res_col9.pdf.
- Leedy, P.D. & Omrod, J.E. (2010). *Practical Research: Planning and Design 9th Ed.* New Jersey: Pearson education, Inc.
- Liebowitz, B. (2009). What's inside the suitcases? An investigation into the powerful resources students and lecturers bring to teaching and learning. *Higher Education Research and Development*, 28(3), 261-274.
- Merlin, G. (2000). Pragmatism and Self-Organization Research Collaboration on the individual level, *Research Policy*, 29(1), 31-40.
- Narvaez-Berthelemot, N., Russell, J. M., Arvanitis, R., Waast, R., & Gaillard, J. (2001). Science in Africa: An overview of mainstream scientific output. In: M. Davis & C. S. Wilson (eds.). *Proc. 8th International Conference on Scientometrics and Informetrics*, Sydney, July, 16-20, 2, 469-476.
- Neuman, W.L. (2006). *Social Research Methods: qualitative and quantitative approaches*. (6th ed.) Boston: Allyn and Bacon.
- Ocholla, D N. (2008). The current status and challenges of collaboration in library and information studies (LIS) education and training in Africa". *New Library World*, 109(9/10), 466 – 479.
- Ocholla, D.N. & Bothma T.D.J. (2007). Library and information education and training in South Africa. *New Library World*, 108(1/2), 55-78.
- Onyancha, O.B. (2009). Towards Global Partnerships in Research in sub-Sahara Africa: an informetric study of the national, regional and international country collaboration in HIV/AIDS literature in eastern and southern Africa. *South African Journal of Libraries and Information Science*, 75(1), 86-99.
- Onyancha, O.B. & Maluleka, J.R. (2011). Knowledge production through collaborative research in sub-Saharan Africa: how much do countries contribute to each other's knowledge output and citation impact? *Scientometrics*, 87, 315–336.
- Onyancha, O.B. & Ocholla, D.N. (2007). Country-wise collaborations in HIV/AIDS research in Kenya and South Africa, 1980–2005. *Libri*, 57(4), 239-254.
- Smith, D. & Katz, S. (2000). Collaborative Approaches to Research. A report to the Higher Education

- Funding Council for England. Centre for Policy Studies in Education, University of Leeds.
- Sooryamoorthy, R. (2009). Do types of collaboration change citation? Collaboration and citation patterns of South African science publications. *Scientometrics*, 81(1), 177-193.
- Sugimoto, C.R. (2010). Collaboration in information and library science doctoral education. *Library & Information Science Research*, 33, 3-11.
- Sun, Y. (2006). Bibliometric analysis of scientific research collaboration between Japan and China. *Int. Workshop on Webometrics, Informetrics and Scientometrics & Seventh COLLNET Meeting*.
- Yamashita, Y. & Okubu, Y. (2006). Patterns of scientific collaboration between Japan and France: Inter-sectoral analysis using Probabilistic Partnership Index (PPI). *Scientometrics*, 68(2), 303-324.

The Diffusion of Nanotechnology Knowledge in Turkey

Hamid Darvish¹ and Yaşar Tonta²

¹*hssdarvish@gmail.com*

Kastamonu University, Faculty of Arts & Sciences, Information Management Department, Kastamonu (Turkey)

²*yasartonta@gmail.com*

Hacettepe University Department of Information Management, 06810 Beytepe, Ankara (Turkey)

Abstract

This paper assesses the diffusion of nanoscience and nanotechnology in Turkey in the last decade using bibliometric and Social Network Analysis (SNA) techniques. We extracted a total of 10,062 articles and reviews from Web of Science (WoS) authored by the Turkish scientists between 2000 and 2011. We divided the data set into two 6-year periods (2000-2005 and 2006-2011). Almost three quarters (7,398) of all papers were published between 2006 and 2011. For each period, we compared the number of nanotechnology papers, the universities' output along with their levels of collaboration with one another, the diffusion and adoption of nanotechnology, the most prolific authors and the nanotechnology research topics studied most often by the Turkish researchers. We found that nanotechnology research and development (R&D) in Turkey is on the rise and its diffusion and adoption has increased tremendously in the second period. This is due primarily to the fact that the government identified nanotechnology as a strategic field a decade ago and decided to provide constant support for nanotechnology R&D. Overlay maps showed that nanotechnology R&D in Turkey concentrated primarily in Materials Sciences, followed by Chemistry, Physics, Clinical Medicine and Biomedical Sciences.

Conference Topics

Country-level studies, Mapping and visualization

Introduction¹

Nanoscience and Nanotechnology is the study of materials at atomic levels within the 1 to 100 nm range (i.e., at a magnitude of 10^{-9} of a meter) (Mehta, 2002). Although Nanotechnology has been introduced more than half a century ago by Feynman (1960), it took some time for the nanotechnology research to pick up. Many countries have invested heavily in nano-related technologies in the past two decades. The US government, for example, has allocated 1.74 billion US dollars to nano-related technologies in 2011 (Sargent, Jr., 2013). European countries under the 7th Framework Program have also heavily invested in joint projects among its members. Consequently, the number of scholarly publications in nano-related technologies in North America, Europe and Far Eastern countries has increased. Turkey as a developed country prepared its strategic plan by taking nano-related research and development into account. Nanotechnology including nanophotonics, nanoelectronics, and nanoscale quantum computing is one of the eight strategic fields of research and technology mentioned in Turkey's "Vision 2023 Technology Foresight Study" that was prepared as part of the "National Science and Technology Policies 2003-2023 Strategy Document" by the Supreme Council of Science and Technology (SCST) more than a decade ago (Ulusal, 2004, pp. 19-20). Nanotechnology as a research field has been receiving state support since 2007 in Turkey (about one billion Turkish Lira, or circa 500 million USD). The Turkish Scientific and Technological Research Council (TUBITAK) and the Ministry of Development (MoD) support nanotechnology projects financially. For example, MoD continues for more than a decade to invest to improve the infrastructure of nanotechnology research facilities and

¹ This paper is based on the findings of first author's Ph.D. dissertation entitled "Assessing the diffusion of nanotechnology in Turkey: A Social Network Analysis approach." (Darvish, 2014).

supported the establishment of nanotechnology research centers. In addition, it supports several nanotechnology-related projects carried out by research institutes and universities.

Thanks to state support, nanotechnology has become a major field of research in Turkey. Universities invested heavily in nanotechnology in the last decade. More than 20 nanotechnology research centers were set up mostly in universities. Among them are Bilkent, Middle East Technical, Hacettepe, Sabancı, İstanbul Technical and Boğaziçi Universities. More than 10 universities are offering both undergraduate and graduate degrees in nanotechnology. More than 100 commercial companies and start-ups of various sizes have also invested in nanotechnology (e.g., Normtest, Arçelik, Yaşar Holding, Yeşim Textile and Zorlu Energy) and developed commercial nanotechnology products in a number of sectors including surface coating, textile, chemistry, automotive and construction industries, and polymer and composite materials. Turkey has been among the first three countries in terms of the growth of nanotechnology research with some 2,000 scientists working in this field (Bozkurt, 2015, p. 49; Denkbaş, 2015, p. 84; Özgüz, 2013). The number of nanotechnology related scientific papers published by Turkish researchers and listed in Web of Science (WoS) is ever increasing (more than 2,500 in 2014 alone).²

This paper aims to assess the diffusion of nanoscience and nanotechnology in Turkey between 2000 and 2011 using bibliometrics and Social Network Analysis (SNA) techniques. It identifies the total production of nano-related publications by Turkish researchers and the key fields in which nanotechnology is applied in Turkey (e.g., biomedicine, pharmacy, and metallurgy). The adoption of nanotechnology by the most prolific universities and the diffusion of nanotechnology knowledge through collaboration among them is also studied.

Literature Review

Scientists have investigated the diffusion of innovation and knowledge in societies from different perspectives. Rogers (2003, p. 5) defines the diffusion of an innovation as “the process by which an innovation is communicated through certain channels over time among the members of a social system.” Social interactions between scientific domains and practitioners are instrumental to the diffusion of innovation and knowledge. According to Rogers, the key elements in the diffusion process are: innovation/knowledge, communication channels, time and social systems (p. 7). An innovation starts with a few people and has a few adopters, but eventually it gains the momentum until it reaches its peak. Rogers likens the diffusion process of an innovation to a mathematically-based bell curve (also known as “Rogers adoption/innovation curve”) and categorizes the adopters accordingly (i.e., starting from the left tail of the curve to the right, 2.5% of the adopters are called “innovators”, 13.5% “early adopters”, 34% “early majority”, 34% “late majority”, and the remaining 16% on the right tail of the curve as “laggards”).

Poire (2011) looks at the timeframe of the adoption of innovations along with the impact of innovations on the economy. He argues that “it takes about 28 years for a new technology to become widely accepted, followed by a period of rapid growth lasting about 56 years. Some 112 years after invention, the innovation reaches maturity and grows in-line with population increases” (Roy, 2005, p. 9). Using these yardsticks, he convincingly charted the adoption processes of textiles, railways, automobiles, computers and nanotechnology. He predicts that nanotechnology, which according to him came into being in 1997, will be more widely adopted by 2025, followed by a 56-year long rapid adoption period (until 2081) during which time nanotechnology products will become an integral part of our everyday life like computers.

² Search on WoS was carried out on January 11, 2015.

If an innovation is communicated among the members of a social system, as Rogers indicated, then studying social systems is important because scientists work and collaborate within such systems. Assessing social relations among scientists reveals how collaborative they are. Conventionally, Derek de Solla Price (1965) studied the scholarly communication process between scientists, thereby opening the door to the quantitative study of science.

Social Network Analysis is a paradigm in which relational interaction among members signifies the role of people in a network structure (Wellman & Berkowitz, 1997). The diffusion of knowledge in a network of people can thus be studied by exploring the social structure of the network along with the relations and collaboration (or lack thereof) among network members using SNA concepts such as density and centrality. For example, poorly connected “structural holes” in a densely connected network are crucial for connecting “clusters” (groups of people) in a network structure and for the diffusion of knowledge in the network (Burt, 1992). Newman (2000) referred to clustering as “community structure”. The value of a person in a social network is therefore linked to his/her potential to establish connections between clusters that are separated by structural holes.

Scientific discovery comes with a group of specialized people who “attend, read and cite the same body of literature and attend the same conferences” (Chen et al., 2009, p. 192). Bibliometric methods such as co-citation (Crane, 1972) or co-author (Girman & Newman, 2002) analyses were used to study the diffusion of knowledge in the network of scientists as well as to track the level of collaboration among different partners along with the emergence of new research areas. As a collaborative model involving universities (research centers), funders and industries, the Triple Helix was proposed to streamline the diffusion of knowledge (Leydesdorff & Etzkowitz, 1998).

Scientometricians use visualizations in addition to other indicators to track or investigate new scientific developments over time. For example, science overlay maps were introduced as a novel approach to illustrate the bodies of research precisely surrounded by global scientific domains (Rafols, Porter & Leydesdorff, 2010). Science overlay maps can represent different types of data and large data sets such as network of authors, publications and universities succinctly and “help benchmark, explore collaborations, and track temporal changes” (Rafols, Porter & Leydesdorff, 2010, p. 1871).

Nanotechnology has been the subject of several studies in the past and reviewing them is beyond the scope of this paper. However, we should mention Milojević (2009, 2012) who studied the cognitive content of nanoscience and nanotechnology as well as its diffusion using SNA techniques and mapped the evolution and socio-cognitive structure of it. We should also mention one particular study that measured the growth and diffusion of nanotechnology on a global level on the basis of the number of publications produced by countries as well as the most prolific institutions and authors along with the most cited authors, papers and journals (Kostoff, Stump, Johnson, Murday, Lau & Tolls, 2006). China, Far Eastern countries, USA, Germany and France were among the most prolific ones.

As mentioned earlier, Turkey is among the first three countries based on the growth of nanotechnology research. Turkey’s contribution to nanotechnology literature was also evident at the global level (Kostoff, Koytcheff & Lau, 2007). Recently, the state of nanotechnology centers and companies carrying out research and manufacturing nano-related technologies in Turkey was studied with a view to compare them quantitatively with their counterparts in China, India and Germany, for example (Aydoğan-Duda & Şener, 2010; Aydoğan-Duda, 2012). The present study attempts for the first time to map the nanotechnology output of Turkish universities and investigate the diffusion of nanoscience and nanotechnology knowledge in Turkey at the micro level by means of Social Network Analysis and bibliometrics. The results can be considered as a stepping stone for comparative studies for future studies.

Method

The aim of this research is to assess the diffusion of nano-related technology by mapping of collaborative social structure of scientists in Turkey between 2000 and 2011. We attempted to address the following issues: (a) the most prolific universities publishing nanotechnology research; (b) the rate of diffusion of nanotechnology knowledge and its adoption within universities between 2000 and 2011; and (c) key areas of nanotechnology research.

In order to answer the research issues, we used a compound textual query on nanotechnology modified from Kostoff's³ and searched (WoS). We retrieved a total of 10,062 papers (with at least one author of each paper affiliated with a Turkish university or research institute) published between 2000 and 2011. We then divided the data set into two 6-year periods (2000- 2005 and 2006-2011) to further assess the diffusion of nano-related technology in Turkey.

We analyzed co-occurrences among universities to capture collaborations in network structures. VOSviewer was used to implement the method of “associative strength” that clustered bibliometric data based on their similarities and mapped the network structure. A geocoder⁴ was used to get the geo-coordinates for each city and Google Maps was used to overlay the relationships among cities on a geographic map. Bibexcel was used to calculate the most frequent collaborators from selected universities in the research. The top ranked universities in each period (2000-2005 and 2006-2011) were selected on the basis of their co-occurrence in terms of scientific collaboration on nanotechnology. Gephi, VOSviewer and GoogleMaps were used to map the network structure.

Findings

The number of Turkey's scientific publications on nanotechnology increased from 215 papers in 2000 to 1,748 in 2011, more than an eight-fold increase (Fig. 1). Almost three quarters (7,398) of all papers (articles and reviews) were published between 2006 and 2011 while the rest (2,664) were between 2000 and 2005. This increase is mainly due to Turkey's making nanotechnology a priority field in its 2003-2023 strategic plan and providing state support to nanotechnology research and development starting from 2007. The number of newly-established universities, hence the number of researchers studying nanotechnology, has also increased tremendously in this period.

There are about 180 universities in Turkey, two-thirds being state-funded. Using the fractional counting method, Figure 2 shows the top ranked universities based on the number of nanotechnology papers they published between 2000 and 2011. The Middle East Technical, Hacettepe, İstanbul Technical, Gazi and Bilkent Universities are the top ranking ones. All but four (Bilkent, Koç, Fatih and Sabancı) universities in Figure 2 are state funded.

³ Personal communication with Prof. Ronald N. Kostoff (20 April 2012). Search query is available from the authors upon request.

⁴ Available from <http://www.gpsvisualizer.com/geocoder/>.

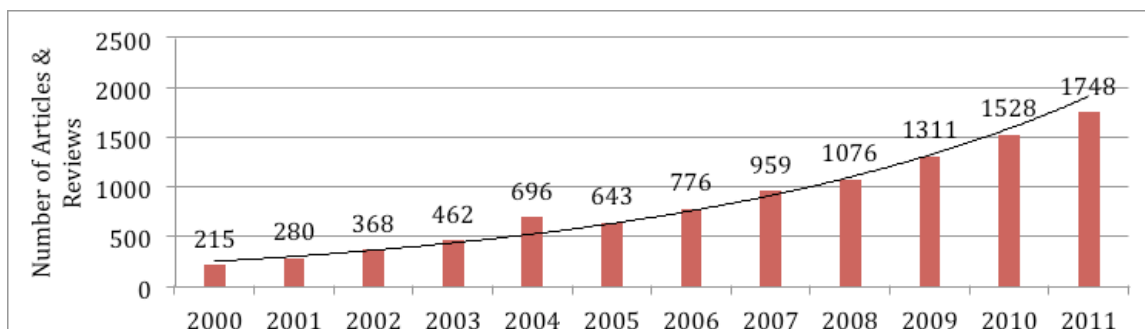


Figure 1. Number of nano-related technologies publications in Turkey: 2000-2011 Source: Thomson's ISI Web of Science as of November 2013.

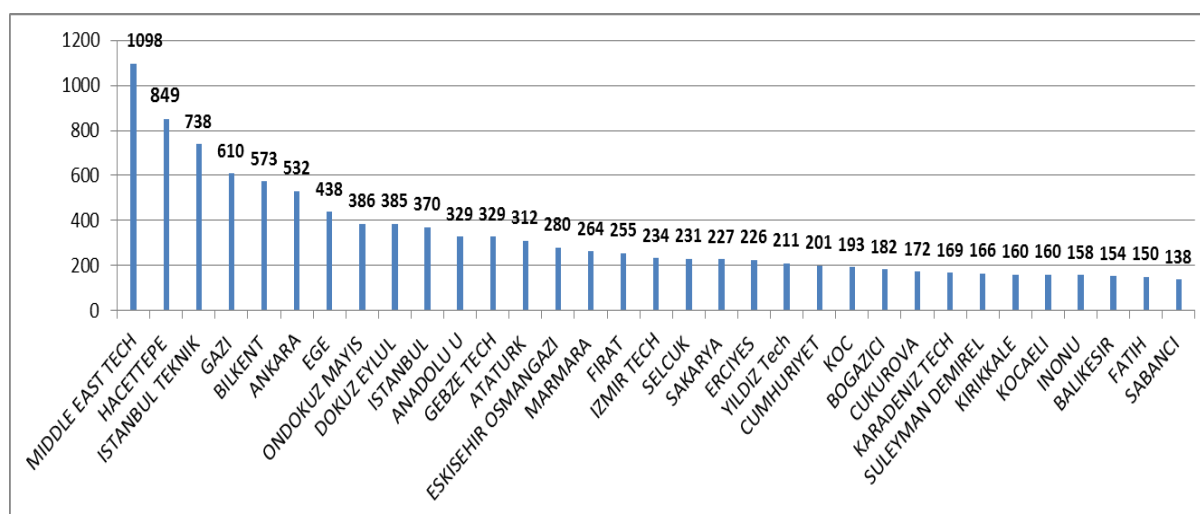


Figure 2. Number of nanotechnology papers of the top Turkish universities between 2000 and 2011 Source: Web of Science as of November 2013.

To assess the level of collaboration and the diffusion of nanotechnology knowledge among universities, we examined the average co-occurrence frequencies of all universities in published papers and created separate networks for the periods of 2000-2005 and 2006-2011 (Fig. 3). The collaboration network was much sparser in the first period with a few universities such as Hacettepe and METU acting as hubs of research on nanotechnology and cooperating with others. The network was much denser in the second period with more universities both acting as hubs of nanotechnology research and collaborating with their counterparts. This is an indication of an increasing level of collaboration among universities in carrying out nanotechnology research within a relatively short period of time.

The diffusion of nanotechnology knowledge in Turkey can be examined from a somewhat different angle by looking at the number of provinces where nanotechnology research took place. Turkey is divided into 81 administrative provinces. The information presented in Figure 4 is less granular than that in Figure 3 due to a few provinces such as İstanbul, Ankara and İzmir having several universities (both old and new). Nevertheless, the number of provinces where nanotechnology research is carried out went up from 40 in the first period (2000-2005) to 72 in the second period (2006-2011). The geographical spread is due to new universities being established in some provinces for the first time and to the government support that enabled researchers both in new and old universities to collaborate further.

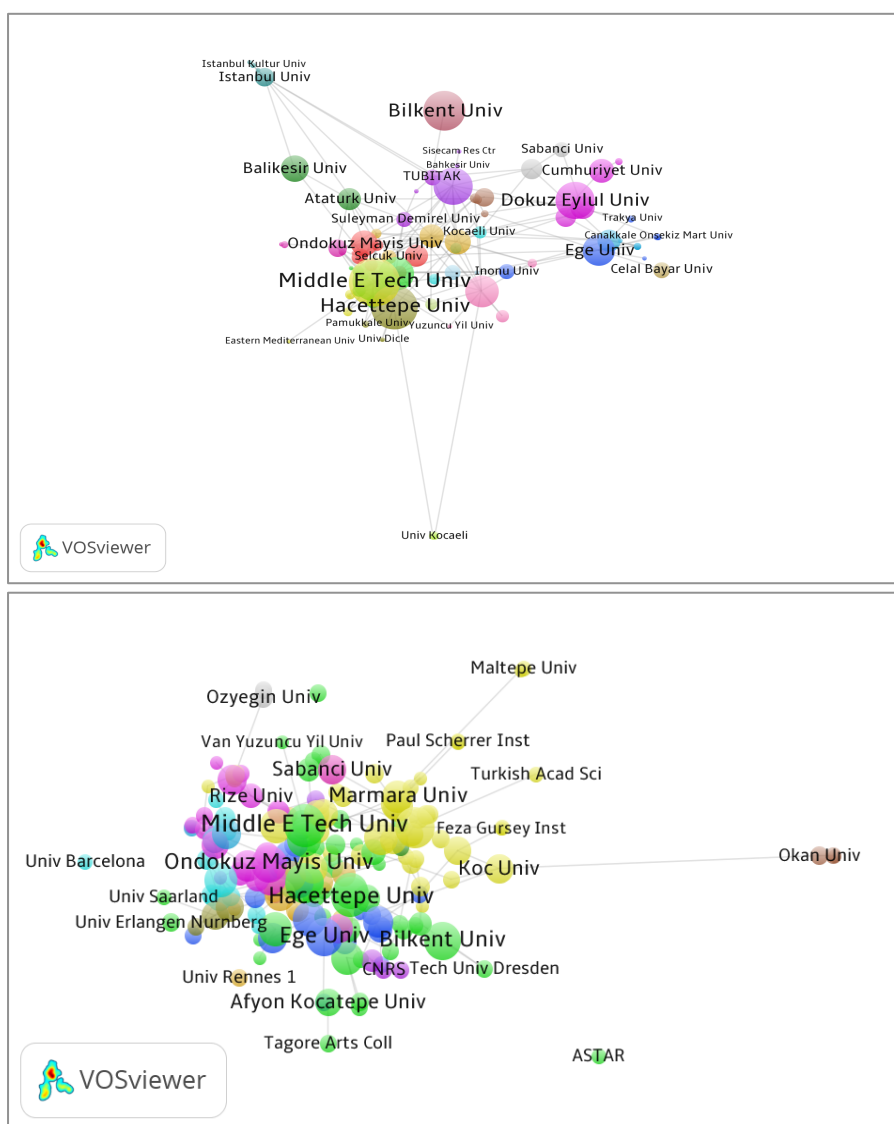


Figure 3. Collaboration of Turkish universities on nanotechnology (top) 2000-2005; (bottom) 2006-2011.

Table 1 shows the top 15 universities with the highest co-occurrence frequencies in both periods. The average co-occurrence frequency for the top 15 universities has almost tripled from 17 in 2000-2005 to 46 in 2006-2011. Note that the top 15 universities in the second period are slightly different from the ones in the first period, as some of the more prolific and more collaborative universities with higher frequencies of co-occurrence replaced the previous ones. We used the fractional counting method and found that the average number of nanotechnology papers published by the top 15 universities in the first period increased from 9 in 2000 to 27 in 2005, and from 35 in 2006 to 77 in 2011 in the second period, indicating more than an eight-fold increase (Table 2).

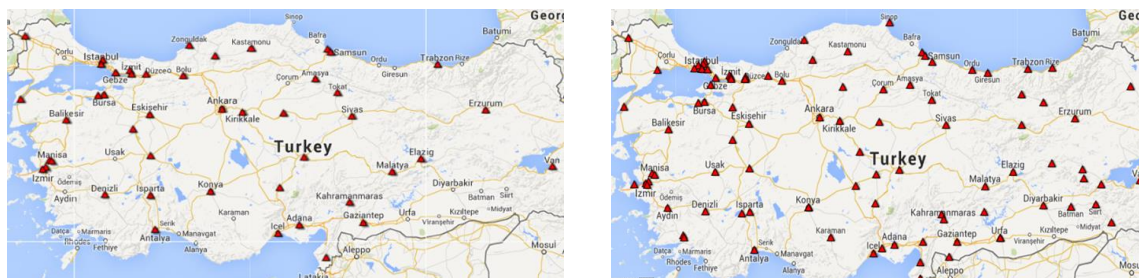


Figure 4. Geographical distribution of nanotechnology research activities in Turkish provinces (l) 2000-2005; (r) 2006-2011.

Table 1. Top 15 Turkish universities with the highest co-occurrence frequencies of collaboration between 2000 and 2011

2000-2005		2006-2011	
University	N	University	N
Hacettepe	30	Hacettepe	63
Middle East Technical	29	Gazi	63
Ankara	21	Middle East Technical	60
Gazi	20	Istanbul Technical	57
Istanbul Technical	18	Ankara	53
Gebze Institute of Technology	17	Gebze Institute of Technology	47
Dokuz Eylül	15	Ondokuz Mayıs	42
Marmara	14	Ege	41
Bilkent	14	Istanbul	41
Abant İzzet Baysal	13	Erciyes	40
Kırıkkale	12	Bilkent	38
Ege	12	Dokuz Eylül	34
Ondokuz Mayıs	11	Anadolu	34
Erciyes	11	Atatürk	33
Kocaeli	11	Fırat	31
Average	17	Average	46

Table 2. Number of papers published by universities with the highest co-occurrence frequencies in the second period (2006-2011)

University	2006	2007	2008	2009	2010	2011
Hacettepe	79	85	89	97	95	107
Gazi	36	77	95	85	99	98
Middle East Technical	77	93	59	131	143	143
İstanbul Technical	52	64	65	88	91	121
Ankara	40	62	70	49	73	54
Gebze Institute of Technology	20	25	33	45	49	55
Ondokuz Mayıs	37	32	35	55	76	74
Ege	16	39	28	60	95	77
İstanbul	25	28	30	42	57	63
Erciyes	16	12	20	41	32	45
Bilkent	34	41	58	63	61	99
Dokuz Eylül	31	43	35	51	52	58
Anadolu	15	29	39	41	45	55
Atatürk	23	18	37	33	55	53
Fırat	17	19	23	31	45	50
Average	35	44	48	61	71	77

Next, we examined the diffusion of nanotechnology knowledge in Turkey using a more refined approach and identified the new authors collaborating each year in order to find out the adoption rate of nanotechnology research. Regardless of whether they appeared in the same paper or not, each new collaboration between any two authors was counted as one and considered a new adoption. The number of collaborating authors was just 214 at the beginning (2000) whereas it rose to 2,989 in 2011 (Table 3 and Figure 5). The number of new adopters was rather slow in the first period (2000-2005) with an average of 216 collaborations per year but the “tipping point” seems to have been reached in 2006 when the number of new adopters jumped from 282 in 2005 to 1622, an almost six-fold increase. The average number of new adopters in the second period (2006-2011) rose to 1868, more than eight times of what it was in the first period. Altogether, the number of cumulative new adopters soared in 12 years and was 13,692 in 2011. The annual rate of cumulative increase in percentages ranged between 11% (2004) and 54% (2006). Needless to say, the increase in the number of new adopters is primarily due to nanotechnology becoming a major research field in Turkey and nanotechnology research being supported by government funds.

Table 3. Number of new and cumulative adopters between 2000 and 2011

Year	# of new adopters	# of cumulative adopters	Rate of cumulative increase (%)
2000	214	214	0
2001	177	391	45
2002	193	584	33
2003	381	965	39
2004	115	1080	11
2005	282	1362	21
2006	1622	2948	54
2007	1668	4652	37
2008	1907	6559	29
2009	1919	8478	23
2010	2225	10703	21
2011	2989	13692	22

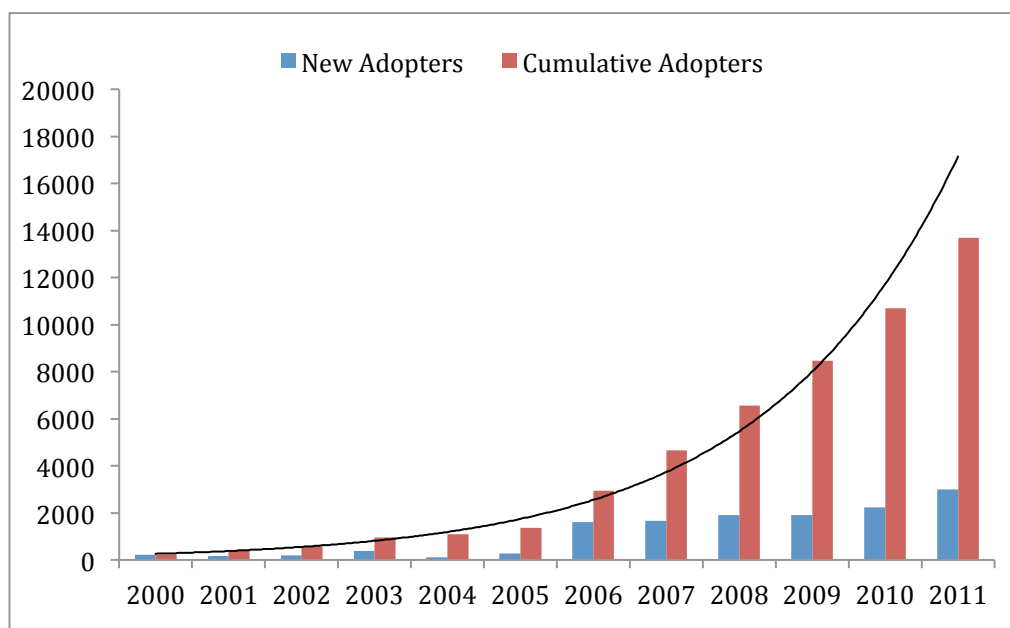


Figure 5. The growth of adoption of nanotechnology knowledge based on the number of collaborating authors (2000-2011).

Next, we identified the most prolific Turkish researchers in nanotechnology between 2000 and 2011 based on the number of papers they authored or co-authored. The fractional counting method was used for co-authored papers. Table 4 shows the top 20 researchers in both periods along with their total number of co-authors. The total number of papers authored or co-authored by the top 20 researchers almost doubled in the second period (from 645 to 1,189). Nine researchers appeared in both periods (italicized in the table) with different ranks. This means that 11 new researchers became more productive than they were in the first period and replaced the less productive ones in the second period or they entered the field anew. O. Buyukgungor of Ondokuz Mayıs University, for instance, is at the top of the second period with 149 papers to his credit even though he did not appear in the top 20 of the first period. The top 20 most prolific researchers co-authored more papers with their colleagues in the second period (216 and 315, respectively). The number of co-authors of nine researchers who appeared in both periods increased 42% in the second period, indicating that they were influential in diffusing the nanotechnology knowledge to their colleagues. The same can probably be said for the remaining 11 researchers who appeared in the top 20 list in the second period.

Finally, we identified the research topics in nanotechnology that were studied more often by the Turkish scientists. We created separate overlay maps of research topics for both periods using ISI's 224 Subject Categories listed in WoS records. Both co-authorship networks and overlay maps were shared with five senior and five junior experts in nanoscience whose publications appeared in leading journals and their comments with respect to their places in the network were recorded (not reported here) (Darvish, 2014).

Table 4. The most prolific Turkish scholars in nanotechnology (2000-2011) Source: WoS (as of November 2013)

2000-20005			2006-2011		
N	First author & affiliation	# of co- authors	N	First author & affiliation	# of co-authors
53	Erkoc S (METU)	29	149	Buyukgungor O (Ondokuz Mayıs)	37
49	<i>Sokmen I (Dokuz Eylül)</i>	16	78	<i>Yagci Y (ITU)</i>	19
42	<i>Ciraci S (Bilkent)</i>	13	75	<i>Denizli A (Hacettepe)</i>	18
39	<i>Denizli A (Hacettepe)</i>	12	72	Yakuphanoglu F (Firat)	28
38	<i>Yagci Y (ITU)</i>	10	67	Ozkar S (METU)	23
37	Celik E (Bilkent)	11	67	<i>Toppare L (METU)</i>	15
37	<i>Sari H (Bilkent)</i>	11	64	<i>Ozbay E (Bilkent)</i>	13
36	Turker L (METU)	28	62	Yesilel OZ (Osmangazi)	17
30	<i>Yilmaz VT (Dokuz Eylül)</i>	8	61	<i>Sokmen I (Dokuz Eylül)</i>	17
30	<i>Toppare L (METU)</i>	7	58	Ozcelik S (Gazi)	12
29	Hascicek YS (Gazi)	8	52	Demir HV (Bilkent)	13
28	Ovecoglu ML (ITU)	7	49	Baykal A (Bilkent)	10
27	Elmali A (Ankara)	8	45	Turan R (METU)	10
26	Elerman Y (Ankara)	8	44	Sahin E (Bilkent)	11
26	Piskin E (Hacettepe)	8	44	<i>Yilmaz VT (Dokuz Eylül)</i>	13
26	<i>Kasapoglu E (Cumhuriyet)</i>	8	43	Caykara T (Gazi)	15
26	Balkan N (Bilkent)	5	41	<i>Sari H (Ankara)</i>	9
22	Yilmaz F (METU)	6	40	<i>Ciraci S (Bilkent)</i>	12
22	Turan S (Marmara)	8	39	<i>Kasapoglu E (Cumhuriyet)</i>	12
22	<i>Ozbay E (Bilkent)</i>	5	39	Albayrak C (Ondokuz Mayıs)	11

Each color in the map represents a subject category and the node size is proportional to its co-occurrence frequency with other nodes (Fig. 6). It appears that the nanotechnology papers authored by Turkish researchers in both periods were primarily related with Materials Science

(black) followed by Chemistry (blue), Physics (purple), Clinical Medicine (red), Biomedical Sciences (light green), Environmental Science and Technology (orange), and Computer Science (fuchsia). Subject categories appeared in overlay maps clearly show the priorities of Turkey in nanotechnology research and development and are commensurate with the nanotechnology products developed by commercial companies based in Turkey.

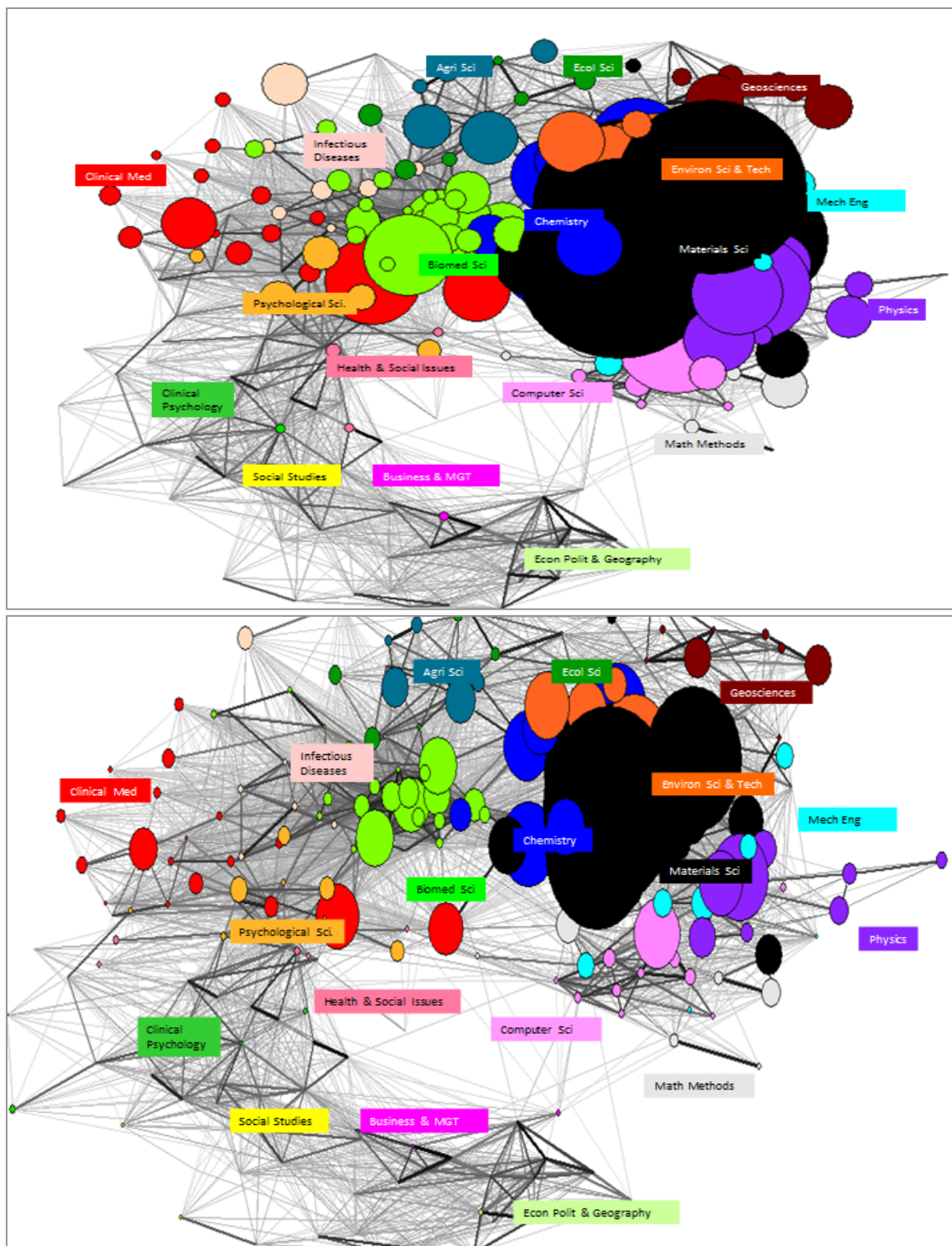


Figure 6. Overlay maps of subject categories of nanotechnology papers authored by Turkish scientists (top) 2000-2005; (bottom) 2006-2011.

Conclusion

Our analysis clearly shows that nanotechnology R&D in Turkey is flourishing. The number of nanotechnology papers published by Turkish scientists has tripled once the Turkish government has identified nanotechnology as one of the eight strategic fields in its national science and technology policy of 2003-2023 and decided to invest in nanotechnology accordingly. This decision has tremendously increased the diffusion and adoption of nanotechnology as a research field. Nanoscientists became more collaborative and more prolific in their research. This is somewhat similar to the experience of India, China, Iran and Latin American countries in that the importance of nanotechnology has increased once they identified it as a promising technology in their national development plans (Aydoğan-Duda, 2012).

The key areas of nanotechnology research and applications in Turkey are primarily in Materials Science, Chemistry, Physics, Clinical Medicine and Biomedical Sciences. All but Clinical Medicine appear in Milojević's list of areas as having the highest number of nanoscience and nanotechnology papers published in the literature (Milojević, 2012). The diversity of nanotechnology research shows that Turkish scientists are well aware of the trans- and interdisciplinary nature of nanotechnology as a discipline, although collaborative nanotechnology research in some areas such as Mathematics, Computer Science and Social Sciences seems to be currently lacking in Turkey.

Nanoscience stimulates scientific research in Physics, Chemistry, Biology and Medicine. Results revealed that notably well-established universities are instrumental in nanoscience research and newer universities are catching up. Turkey recognized the importance of nanotechnology as a strategic field relatively early. Based on Poire's timeframe of innovations becoming the drivers of economy, we can say that the diffusion of nanotechnology and its widespread adoption in Turkey will likely continue to accelerate until early 2030s.

References

- Aydoğan-Duda, N. (2012). Nanotechnology: A Descriptive Account. Making it to the Forefront in Aydoğan-Duda, N. (Ed). *Nanotechnology: A Developing Country Perspective*. 1, (pp. 1-4). New York: Springer.
- Aydoğan-Duda, N., & Şener, I. (2010). Entry Barriers to the Nanotechnology Industry in Turkey in Ekekwe, N. (Ed). *Nanotechnology and Microelectronics: Global Diffusion, Economics and Policy*. (pp. 167-173). Hershey, PA: IGI Global.
- Bozkurt, A. (2015). Türkiye, 10 yıldır “en küçük” dünyanın farkında, artık büyük adımlar atması gerekiyor (Turkey is aware of the “smallest” world for 10 years, but it should take big steps). *Bilişim: Aylık Bilişim Kültürü Dergisi*, 43(172), 44-53. <http://www.bilisimdergisi.org/>
- Burt, R.S. (1992). *Structural Holes: The Social Structure of Competition*, Cambridge, MA: Harvard University Press.
- Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z., & Pellegrino, D. (2009). Towards an Explanatory and Computational Theory of Scientific Discovery. *Journal of Informetrics*, 3(3), 191-209.
- Crane, D. (1972). *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. Chicago, IL: University of Chicago Press.
- Darvish, H. (2014). *Assessing the Diffusion of Nanotechnology in Turkey: A Social Network Analysis Approach*. Unpublished PhD Dissertation, Hacettepe University, Ankara.
- Denkbaş, E.B. (2015). Nanoteknolojiye yapılacak yatırımlar, ülkelerin ekonomik gücünü yansıtabilecek bir parametre olacak (Investments in nanotechnology will become a parameter reflecting economic powers of countries). *Bilişim: Aylık Bilişim Kültürü Dergisi*, 43(172), 78-87. <http://www.bilisimdergisi.org/>
- Feynman, R.P. (1960). There's Plenty of Room at the Bottom. Caltech Engineering & Science. Retrieved, Feb.14, 2014, from <http://calteches.library.caltech.edu/1976/1/1960Bottom.pdf>
- Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks, *PNAS*, 99(12), 7821–7826.
- Kostoff, R.N., Koytcheff, R.G., & Lau, C.G.Y. (2007). Global nanotechnology research literature overview, *Technological Forecasting & Social Change*, 74, 1733-1747.

- Kostoff, R.N., Stump, J. A., Johnson, D., Murday, J.S., Lau, C.G.Y., & Tolls, W.M. (2006). The structure and infrastructure of global nanotechnology literature. *Journal of Nanoparticles Research*, 8, 301-321.
- Leydesdorff, L. & Etzkowitz, H. (1998). The Triple Helix as a model for innovation studies (Conference Report), Science & Public Policy, *Research Policy* 25(3), 195-203.
- Mehta, M., (2002). Nanoscience and nanotechnology: Assessing the nature of innovation in these fields. *Bulletin of Science, Technology and Society*, 22(4), 269-273.
- Milojević, S. (2012). Multidisciplinary cognitive content of nanoscience and nanotechnology. *Journal of Nanoparticle Research*, 14(1), 1-28.
- Milojević, S. (2009). *Big Science, Nano Science? Mapping the Evolution and Socio-Cognitive Structure of Nanoscience/Nanotechnology Using Mixed Methods*. Unpublished PhD Dissertation, University of California, Los Angeles.
- Newman, M. E. J. (2000). The structure of scientific collaboration networks. *PNAS*, 98(2), 404–409.
- Özgüz, V. (2013). Nanotechnology Research and Education in Turkey (presentation slides). Retrieved, December 27, 2014, from: http://rp7.ffg.at/upload/medialibrary/12_Oezguez.pdf.
- Poire, N.P. (2011). *The great transformation of 2021: How the looming sustainability crisis will revolutionize capitalism, fracture the nation-state, and topple American supremacy*. Lulu.com.
- Price, D. J. de Solla. (1965). Networks of scientific papers. *Science*, 49(3683), 510–515.
- Rafols, I., Porter, A. L., & Leydesdorff, L. (2010). Science overlay maps: a new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, 61(9), 1871–1887.
- Rogers, E. M. (2003). *Diffusion of Innovations*. 5th ed. New York: The Free Press.
- Roy, J.M.A. (2005). Nanotechnology: Investing in products of the future (Research report). Retrieved, January 12, 2015, from <http://www.ianano.org/Hambrecht-report.pdf>.
- Sargent, Jr., J.F. (2013). *Nanotechnology: A policy primer*. Washington, DC: Congressional Research Service. Retrieved, April 28, 2014, from <http://www.fas.org/sgp/crs/misc/RL34511.pdf>.
- Ulusal Bilim ve Teknoloji Politikaları 2003-2023 Strateji Belgesi. (2004, November 2). (National Science and Technology Policies 2003-2023 Strategy Document). Version 19. Turkish Scientific and Technological Research Council. Retrieved, December 6, 2014, from http://www.tubitak.gov.tr/tubitak_content_files/vizyon2023/Vizyon2023_Strateji_Belgesi.pdf
- Wellman, B. & Berkowitz, S. D. (1997). *Social Structures: A Network Approach*. 2nd ed. Cambridge: University of Cambridge.

The Network Structure of Nanotechnology Research Output of Turkey: A Co-authorship and Co-word Analysis Study¹

Hamid Darvish¹ and Yaşar Tonta²

¹*darvish@gmail.com*

Kastamonu University, Faculty of Arts & Sciences, Department of Information Management, Kastamonu (Turkey)

²*yasartonta@gmail.com*

Hacettepe University, Department of Information Management, 06800 Beytepe, Ankara (Turkey)

Abstract

This paper aims to assess the diffusion of nanotechnology knowledge within the Turkish scientific community using co-citation and co-word analysis techniques. We retrieved a total of 10,062 records of nanotechnology papers authored by Turkish researchers between 2000 and 2011 from Web of Science (WoS) and divided the data set into two 6-year periods. We identified the most prolific and collaborative top 15 universities in each period based on their network properties. We then created co-authorship networks of Turkish nanotechnology researchers in each period and identified the most prolific and collaborative top 15 authors on the basis of network centrality coefficients. Finally, we used co-word analysis to identify the major nanotechnology research fields in Turkey on the basis of the co-occurrence of words in the titles of papers. Findings show that nanotechnology research in Turkey continues to increase due to researchers collaborating with their colleagues. Turkish researchers tend to collaborate within their own groups or universities and the overall connectedness of the network is thus low. Their publication and collaboration patterns conform to Lotka's law. They work mainly on nanotechnology applications in Materials Sciences, Chemistry and Physics, among others. This is commensurate, more or less, with the global trends in nanotechnology research and development.

Conference Topic

Country-level studies, Mapping and visualization

Introduction

Nanotechnology is a relatively new field studying materials at atomic levels within the 1 to 100 nanometer (nm) range (one nm is equal to one billionth of a meter, or, 10^{-9}) (Nanotechnology, 2015). It involves physics, chemistry, medicine, and biotechnology, among others, and promises a great deal of innovation for, and benefit to, society as a whole. Turkey identified nanotechnology early on (2003) as one of the eight strategic fields to support and invested considerably in nanotechnology infrastructure and education. It set up several “centers of excellence” in universities for nanotechnology research and development (R&D). Among them are the Research Center for Nanotechnology and Biotechnology of the Middle East Technical University (METU) and the National Nanotechnology Center in Bilkent University. The former is the first such center established with 15M USD government support while the latter is the first largest multi-purpose nanotechnology center established with 70M USD investment. Universities themselves also invested in nanotechnology. Altogether, there are currently more than 20 nanotechnology research centers in Turkey (Bozkurt, 2015;

¹ This paper is based on the findings of first author's PhD dissertation entitled “Assessing the diffusion of nanotechnology in Turkey: A Social Network Analysis approach.” (Darvish, 2014).

Denkbaş, 2015; Özgüz, 2013). The private sector has also invested in nanotechnology in Turkey. Currently, more than 100 companies working in this field and they already developed several nanotechnology products and commoditized them.

In parallel with both government's and private sector's financing of nanotechnology research, several universities initiated multidisciplinary nanotechnology degree programs both at undergraduate and graduate levels (MSc and PhD). The undergraduate and graduate programs of Bilkent University's "Material Science and Nanotechnology", METU's "Micro and Nanotechnology" and Hacettepe University's "Nanotechnology and Nanomedicine" are among them.

The substantial interest and investment in nanotechnology triggered nanotechnology research in Turkey. In fact, Turkey is among the top three countries in the world in terms of the growth rate of nanotechnology research. More than 2,000 researchers are active in this field producing some 2,500 papers in 2014 alone² (Bozkurt, 2015, p. 49; Denkbaş, 2015, p. 84; Özgüz, 2013). In this paper, we investigate the development of nanotechnology research in Turkey using bibliometric and Social Network Analysis (SNA) techniques to study the network characteristics of more than 10,000 papers authored by Turkish researchers between 2000 and 2011. We compare the diffusion of nanotechnology research between 2000-2005 and 2006-2011 by measuring the network properties such as degree, betweenness and closeness centrality coefficients of the most prolific and collaborative universities and researchers for each period. We also identify the major nanotechnology research strands in Turkey using co-word analysis.

Literature Review

Information scientists have studied the growth of science and communication using bibliometrics and Social Network Analysis (SNA). While the former deals mainly with the effects of scientific productivity using citation analysis, the latter mainly focuses on the pattern of relationships among scientists. The network composed of co-authorship among scientists is a true indication of their cooperation in research activity.

The "small world" effect is a phenomenon that has been studied by scientists in different fields. This phenomenon conjectures that each member (node) in a society is linked to others (edges) through friends. Literally, every node in a small world is connected through an acquaintance. Newman (2000) found out that average distance from one person to the other by an acquaintance is proportional to the logarithm of the size of the community, implying one of the small world properties. Moreover, he found out that traversing between the two randomly selected nodes of a network takes an average of six steps.

In social contexts, Moody (2004) analyzed the structure of a social science collaboration network over a period. He discovered that collaboration between graduate students in a specific topic creates a small world of scientists and removes restrictions between them. Small world networks may manifest themselves in several shapes and models. Therefore, a good understanding of small world models helps us understand the network characteristics, too. For example, according to Watts (2003) a social network can be categorized as active or passive. Granovetter (1974) studied an active social network from the perspective of finding a job while Burt (1992) looked at such a network as social capital preluding the "rich get richer" phenomenon. In this study, the co-authorship network of structure is represented in a passive sense where the nodes and the edges connecting them are treated as actors and their relationships. Small world models are comprised of clusters or components. Clusters embedded in a network structure reveal a property called "clustering coefficient". According to Watts and Strogatz (1998), one can define a clustering coefficient C , which is the average

² Search on WoS was carried out on January 11, 2015.

fraction of pairs of neighbors of a node which are also neighbors. That is to say, if node *A* neighbors with node *B* and node *B* is a neighbor of node *C*, then there is a probability that node *A* is also a neighbor of node *C*.

According to Otte and Rousseau (2002, p. 443), betweenness, closeness and degree centrality are well known measures used in analyzing networks. Betweenness centrality is defined as the number of shortest paths going through a node. Thus, a node with high betweenness centrality will have a large impact on the diffusion of knowledge in the network (assuming that knowledge diffusion follows the shortest paths). Centrality is the total number of links that a node has. Degree centrality identifies the most influential node in the diffusion of knowledge in the social network. Closeness measures how far a node is from other nodes in the network structure. Closeness centrality is a measure of how long it will take to diffuse knowledge in a network (Centrality, 2015).

Betweenness centrality plays an important role in the structures of social networks. According to Freeman (2004), the discovery of the structural properties of scientific papers is measured by the betweenness centrality. Actors with a high level of betweenness centrality play a pivotal role in connecting different groups within the network. Betweenness centrality characterizes preferential attachments, cliques, or brokers. Preferential attachments play an important role in network development (Barabasi & Albert, 1999, p. 509). In other words, people in social networks tend to work with well-known people that lead to the concept of “strong and weak ties”, characterizing a group of people attached to one node with high centrality. This is called the “star network model” (Moody, 2004; Scott, 2000).

Newman (2000) stated that collaboration among scientists in networks is a good example of showing preferential attachment. As mentioned earlier, if two nodes have high degrees of centrality, the probability of being acquainted with a mutual friend gets higher. Only a small percentage of people in a social network are well connected while the rest are loosely connected (Lotka’s law). The productivity of authors in a network resembles Lotka’s law in that a small number of researchers publish the majority of papers while large numbers of researchers publish one or two papers (Martin, Ball, Karrer & Newman, 2013). Each group of authors creates a community in which a node with a high degree of centrality is the central node. Therefore, collaboration networks consist of separate clusters representing different scientific fields where they may connect through lower degree connectors. Each community comprises several star networks and these clusters may be connected by a node of lesser degree. Newman (2000) referred to clustering as “community structure”.

Co-authorship analysis is used by bibliometricians to track temporal and topological diffusion of scientific publications. Co-authorship stimulates the knowledge diffusion in scientific communities (Chen et al., 2009, p. 192). Thus, co-authorship analysis is used quite often to study the diffusion of innovation and knowledge. For example, Özel (2010) assessed the diffusion of knowledge in business management among academia in Turkey from 1928 to 2010 by studying the co-authorship relationships of academics in business management.

Co-word analysis of texts helps map scientific fields and reveals the cognitive structure of the scientific domain (Chen, 2004). Callon, Courtial, Turner, and Bauin (1983) used the co-word analysis to study the literature over time in terms of the frequencies or co-occurrences of words in titles, abstracts, or more generally, in text. PageRank measuring the popularity of web pages is a similar metric (Page & Brin, 1989). For example, the appearance of a certain author in the references of a corpus of articles reflects the prestige of that author in the network structure.

As we mentioned earlier, the growth rate of nanotechnology research in Turkey is quite encouraging and researchers contribute to the global nanotechnology literature (Kostoff et al., 2006; Kostoff, Koytcheff & Lau, 2007). Although the state of the art of nanotechnology centers and companies has been studied quantitatively (Aydoğan-Duda & Şener, 2010;

Aydoğan-Duda, 2012), their research output in terms of scientific papers has yet to be studied in detail. This is the first such study to investigate the diffusion of nanotechnology in Turkey and the level of collaboration among the most prolific universities and researchers using co-authorship and co-word analysis.

Method

This paper aims to depict the development of nanotechnology in Turkey between 2000 and 2011 by identifying the network structure of nanotechnology papers authored by Turkish researchers and finding out the most productive universities and researchers who help diffuse the nanotechnology knowledge by collaborating with their peers. Social network analysis, co-authorship and co-word analysis tools were used to map the nanotechnology network structure and the collaboration patterns. We attempt to answer the following research questions:

- 1) Which universities and researchers contribute most to the diffusion of nanotechnology research in Turkey by collaboration?
- 2) Do co-authorship networks in nanotechnology literature exhibit a “small world” network structure?
- 3) What are the main nanotechnology research interests of Turkish scholars?

To answer these questions, we retrieved a total of 10,062 records of nanotechnology papers (articles and reviews) from Web of Science (WoS) published between 2000 and 2011 by Turkish authors. We divided the data set into two equal periods (2000-2005 and 2006-2011) to better identify the trends. Almost three quarters of papers (7,398 papers or 73.5%) were published in the second period. Elsewhere, we presented the descriptive statistics for each period on the number of nanotechnology papers published by universities and analyzed the diffusion and adoption of nanotechnology in Turkey by means of the output of the most prolific authors (Darvish & Tonta, 2015). In this paper, we investigate the diffusion of nanotechnology in Turkey by studying the network properties of nanotechnology literature. We first identified the top 15 most prolific universities and authors by means of social network analysis tools. We then identified the scientists with the highest coefficients of centrality in the network structure. We used co-authorship, co-word³ and factor analyses to track the collaboration patterns and research interests of Turkish nanotechnology scholars between the two periods. We used Bibexcel, VOSviewer, Pajek and Gephi to create files and map the bibliometric data, calculate the properties of the social network structure (e.g., the betweenness, closeness, and degree centralities and the PageRank of each node) and depict the network's features visually.

Findings

Table 1 shows the network properties of the top 15 selected universities in each period (2000-2005 and 2006-2011) ranked by the degree centrality coefficients of their nanotechnology papers. Middle East Technical (METU), Bilkent and Hacettepe Universities are at the pinnacle of the list and they contributed to the network with the highest number of nanotechnology papers. İstanbul Technical (İTÜ), Erciyes and Kocaeli Universities are at the bottom of the list with the lowest degree centrality coefficients in the 2000-2005 period. Nodes with higher degree centralities participate more in the network than that with the lower ones and the network structure adheres to the small world phenomenon.

³The co-word analysis was conducted based on software: <http://www.leydesdorff.net/software/fulltext/index.htm>

Table 1. Centrality coefficients of nanotechnology papers of the top 15 universities between 2000-2005 and 2006-2011

2000-2005					2006-2011				
University	# of papers	Degree centrality	Closeness centrality	Betweenness centrality	University	# of papers	Degree centrality	Closeness centrality	Betweenness centrality
Middle East Technical	353	0.523	0.467	0.113	Bilkent	356	0.620	0.588	0.069
Bilkent	183	0.515	0.495	0.124	Gebze Institute of Technology	227	0.603	0.541	0.068
Hacettepe	283	0.401	0.495	0.072	Hacettepe	552	0.574	0.524	0.022
Ondokuz Mayıs	65	0.357	0.359	0.041	Middle East Technical	646	0.562	0.511	0.054
Dokuz Eylül	108	0.333	0.393	0.109	Istanbul Technical	481	0.534	0.468	0.031
Gebze Institute of Technology	71	0.314	0.499	0.110	Anadolu	224	0.470	0.379	0.042
Kırıkkale	36	0.288	0.457	0.119	Gazi	490	0.457	0.373	0.070
Ege	84	0.276	0.359	0.126	Ondokuz Mayıs	309	0.450	0.415	0.067
Abant İzzet Baysal	11	0.252	0.612	0.184	Istanbul	245	0.445	0.394	0.045
Gazi	127	0.244	0.373	0.156	Ege	315	0.431	0.382	0.035
Marmara	64	0.225	0.336	0.215	Ankara	348	0.418	0.363	0.071
Ankara	181	0.224	0.373	0.072	Dokuz Eylül	270	0.323	0.429	0.060
Kocaeli	21	0.218	0.325	0.425	Firat	185	0.317	0.452	0.051
Erciyes	58	0.162	0.466	0.098	Erciyes	166	0.256	0.452	0.049
Istanbul Technical	214	0.109	0.363	0.151	Atatürk	219	0.230	0.316	0.091
Avg		0.296	0.425	0.141	Avg		0.446	0.439	0.055

The average degree centrality for the top 15 universities rose from 0.296 in the first period to 0.466 in the second period, indicating an almost 60% increase. Istanbul Technical University's degree centrality increased five times between the two periods, making it one of the top nodes in the second period. Kırıkkale, Abant İzzet Baysal, Marmara and Kocaeli Universities with relatively fewer number of papers did not make it to the top 15 universities in the 2006-2011 period and were replaced by Anadolu, İstanbul, Fırat and Atatürk Universities.

Bilkent University is at the top of the 2006-2011 list with the highest closeness centrality coefficient (0.588) followed by Gebze Institute of Technology (0.541) (which was in the 6th place in the first period). Their high closeness centrality coefficients indicate that sub-networks within the whole network are almost 60% connected. However, their betweenness centrality coefficients are relatively low, which means that the flow of information among sub-clusters within the whole network is slow. Hacettepe and Middle East Technical Universities are also at the top of the 2006-2011 list. These four universities form a cohesive network structure in 2006-2011. However, the average closeness centrality coefficient stayed almost the same for both periods (0.425 and 0.439, respectively). In other words, it took equally long to spread nanotechnology knowledge for the top 15 universities in each period.

In general, betweenness centrality coefficients are much lower for all universities. In fact, the average betweenness centrality has decreased from 0.141 to 0.055 in the second period, indicating that sub-clusters in the network structure became less connected in the second period for the top 15 universities. Atatürk, Ankara, Gazi, Bilkent, Gebze Institute of Technology and Ondokuz Mayıs Universities have the highest betweenness centrality coefficients in the second period, an indication of relatively higher flow of information among sub-clusters within the network than the rest. Dokuz Eylül, Hacettepe and Ankara Universities

have the lowest betweenness centrality coefficients in the first period and Hacettepe, İstanbul Technical and Ege Universities in the second period.

Next, we studied the co-authorship network structures in both periods using social network analysis (SNA) techniques (Fig. 1). SNA enabled us to discern the nodes that might be crucial to the diffusion of nanotechnology knowledge. The network consists of 470 nodes and 1,042 edges in 2000-2005 and 945 nodes and 4,915 edges in 2006-2011. The rates of growth for nodes and edges (ties) increased two- and four-folds, respectively, between the two periods. However, the level of collaboration has not changed so much. There is a minimal change in density (from 0.009 to 0.011) between the two periods, but the network is still quite sparse. Nonetheless, the average degree and clustering coefficients show that clusters within the network are somehow connected for both periods. For example, the average clustering coefficient for 2000-2005 is 0.75, indicating that 75% of the nodes were connected. Since the network has grown in the second period, the rate of connectedness has decreased (0.51), indicating that newly formed clusters were not that cohesive yet.

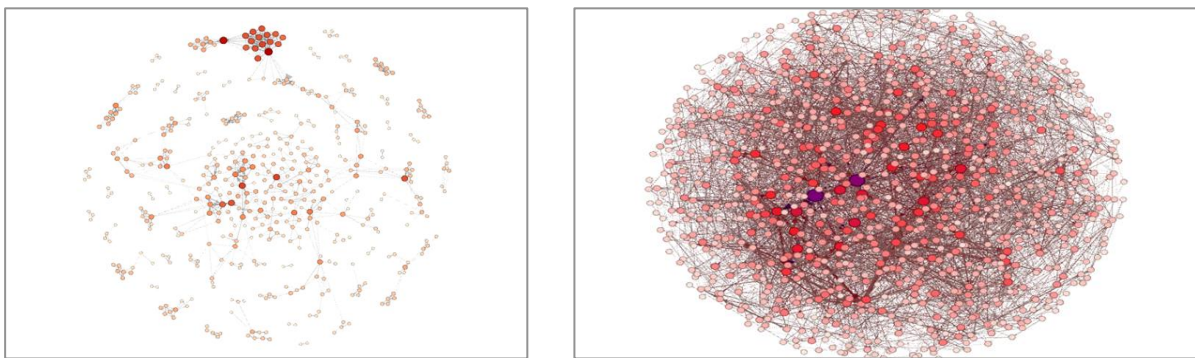


Figure 1. Co-authorship network of scientists working on nanotechnology between: (1) 2005-2011 and (r) 2006-2011

The network in the second period adheres to the transitivity relations, indicating that the network at meso level is well connected even though the sub-clusters are not that well connected (especially in the periphery of the network) (Fig. 1). That is to say that there has been some progress in terms of creating new sub-clusters in the co-authorship network, although links among sub-clusters have yet to be formed. In other words, almost all scientists have co-authored with one or more authors in their own cluster but not beyond.

Table 2 shows the top 15 Turkish authors and their affiliations with the highest centrality coefficients (closeness, betweenness, degree, and PageRank) between 2000 and 2005 who contributed to the diffusion of nanotechnology with their scientific papers. Some scientists appear in more than one columns of centrality due to their high collaboration level in the network structure. For example, Yakuphanoglu F (Firat University), Yağcı Y and Öveçoğlu MN (İTÜ), Çelik E (Dokuz Eylül) and Denizli A (Hacettepe) appeared in three columns with high degree (collaborator), betweenness (broker and gatekeeper), and PageRank coefficients (prolific author) while Yılmaz F and Toppare L (METU), Morkoç H (Atatürk), Özdemir I (Dokuz Eylül) and Pişkin E (Hacettepe) appeared at least in two columns out of four (degree, betweenness, closeness and PageRank centralities). They were highly influential in the diffusion of nanotechnology in Turkey between 2000 and 2005.

Table 2. Network properties of the top 15 Turkish authors based on co-authorship degree centralities: 2000-2005.

Rank	Degree centrality	Betweenness centrality	Closeness centrality	PageRank
1	Balkan N (Fatih)	Yilmaz F (METU)	Sarı H (Bilkent)	Ovecoğlu MN (ITU)
2	Teke A (Balıkesir)	Gencer A (Hacettepe)	Sökmen I (Dokuz Eylül)	Çelik E (Dokuz Eylül)
3	Yağcı Y (ITU)	Koralay H (Firat)	Kasapoğlu E (Cumhuriyet)	Denizli A (Hacettepe)
4	Yakuphanoglu F (Firat)	Okur S (Izmir Inst Tech)	Çiraci S (Bilkent)	Hasçicek YS (Gazi)
5	Ovecoğlu MN (ITU)	Denizli A (Hacettepe)	Aytor O (Bilkent)	Yağcı Y (ITU)
6	Çelik E (Dokuz Eylül)	Yavuz H (Hacettepe)	Biyikli N (METU)	Yakuphanoglu F(Firat)
7	Yilmaz F (METU)	Güneş M (Kırıkkale)	Özbay E (Bilkent)	Toppare L (METU)
8	Toppare L (METU)	Yakuphanoglu F (Firat)	Doğan S (Bilkent)	Yilmaz VT (Ondokuz Mayıs)
9	Doğan S (Bilkent)	Balkan N (Fatih)	Morkoç H (Atatürk)	Pişkin E (Hacettepe)
10	Morkoç H (Atatürk)	Çelik E (Dokuz Eylül)	Sarı B (Gazi)	Erkoç Ş (METU)
11	Denizli A (Hacettepe)	Pişkin E (Hacettepe)	Talu M (Gazi)	Kurt A (Koç)
12	Erol A (Istanbul)	Güven K (Erciyes)	Kartaloğlu (Bilkent)	Elmalı A (Ankara)
13	Özdemir I (Dokuz Eylül)	Yağcı Y (ITU)	Yilgor E (Koç)	Hincal AA (Hacettepe)
14	Turan R (METU)	Ovecoğlu MN (ITU)	Yilgor I (Koç)	Ozdemir I (Dokuz Eylül)
15	Dag O (Bilkent)	Menceloğlu YZ (Sabancı)	Andaç O (Ondokuz Mayıs)	Oral A (Sabancı)

Co-authorship map of the first authors for the first period is shown on the left-hand side of Figure 2. Most of the authors listed in Table 2 are also on the map. Although most authors were from universities with high degree centralities, other authors whose universities did not have high degree centralities were also instrumental in the diffusion of nanotechnology knowledge in the network during the 2000-2005 period (e.g., Yilgor E and Yilgor I from Koç, Koralay H and Yakuphanoglu E from Firat, and Kasapoğlu E from Cumhuriyet Universities).

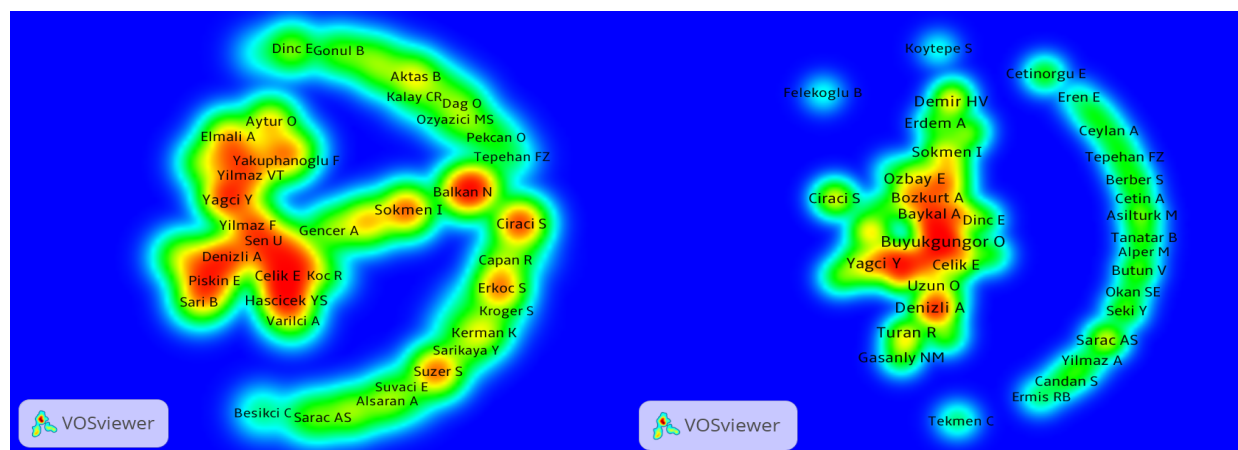


Figure 2. Co-authorship map of Turkish nanotechnology scientists between: (l) 2000-2005 and (r) 2006-2011.

Table 3 shows the top 15 authors who were influential in the diffusion of nanotechnology in Turkey between 2006 and 2011. Interestingly, Büyükgüngör O of Ondokuz Mayıs University has the highest centrality coefficients in all four categories but one (the betweenness centrality) even though he was not in the top 15 authors in the first period. His name appears in the center of the 2006-2011 network of Figure 2 as a prestigious researcher playing an important role in the dissemination of nanotechnology knowledge in the network structure. (His research field is Crystallography.) Similarly, Özçelik S of Gazi University is at the top 15 in all four categories. Six authors appear in at least three columns: Denizli A (Hacettepe), Şahin E (Gazi), Yağcı Y (İTÜ) and Toppare L (METU) in degree, betweenness and PageRank columns, and Özbay E and Çiracı S (Bilkent) in degree, closeness and PageRank columns. An

additional six authors appear in at least two columns: Yeşilel ÖZ (Osmangazi) and Baykal A (Fatih) in closeness and PageRank columns; Yıldız A (Fatih) and Yılmaz F (METU) in degree and betweenness columns; Çakmak M (Koç) in betweenness and PageRank columns; and Turan R (Ege) in degree and PageRank columns.⁴ It should be pointed out that even though Fatih and Karadeniz Technical Universities failed to have the highest degree centrality coefficients in neither period, some of their scientists (e.g., Yıldız A and Bacaksız E, respectively) played an important role nonetheless in the diffusion of nanotechnology knowledge in the network.

The centrality coefficients of four authors were high in both periods: Yağcı Y (İTU), Denizli A (Hacettepe), and Toppare L and Yılmaz F (METU). They were highly active in spreading the nanotechnology knowledge in Turkey between 2000 and 2011 as prolific authors, collaborators, brokers and gatekeepers, and diffusers.

Table 3. Network properties of the top 15 authors based on co-authorship degree centralities: 2006-2011.

Rank	Degree centrality	Betweenness centrality	Closeness centrality	Page Rank
1	Büyükgüngör O (Ondokuz Mayıs)	Yılmaz F (METU)	Büyükgüngör O (Ondokuz Mayıs)	Büyükgüngör O (Ondokuz Mayıs)
2	Şahin E (Gazi)	Büyükgüngör O (Ondokuz Mayıs)	Yeşilel ÖZ (Osmangazi)	Özbay E (Bilkent)
3	Toppare L (METU)	Özçelik S (Gazi)	Demir HV (Bilkent)	Özçelik S (Gazi)
4	Yılmaz F (METU)	Toppare L (METU)	Nizamoglu S (Bilkent)	Toppare L (METU)
5	Özçelik S (Gazi)	Yağcı Y (İTU)	Çağlar Y (Anadolu)	Denizli A (Hacettepe)
6	Yağcı Y (İTU)	Şahin E (Gazi)	İlican S (Anadolu)	Turan R (Ege)
7	Özbay E (Bilkent)	Yıldız A (Fatih)	Çağlar M (Anadolu)	Şahin E (Gazi)
8	Turan R (Ege)	Çakmak M (Koç)	Özbay (Bilkent)	Çıracı S (Bilkent)
9	Çakmak M (Kırıkkale)	Şahin O (Dokuz Eylül)	Özçelik S (Gazi)	Yeşilel ÖZ (Osmangazi)
10	Yerli A (Sakarya)	Yılmaz M (İstanbul)	Baykal A (Fatih)	Yağcı Y (İTU)
11	Yıldız A (Fatih)	Turan R (METU)	Köseoglu Y (Fatih)	Sökmen I (Dokuz Eylül)
12	Çetin K (Ege)	Bacaksız E (Karadeniz Technical)	Toprak MS (Fatih)	Arslan H (Hacettepe)
13	Çıracı S (Bilkent)	Denizli A (Hacettepe)	Çıracı S (Bilkent)	Oskar S (METU)
14	Denizli A (Hacettepe)	Şen S (Yalova)	Durgun E (Bilkent)	Çakmak M (Koç)
15	Sari H (İTU)	Balkan A (Fatih)	Akgöl S (Adnan Menderes)	Baykal A (Fatih)

The collaboration network of Turkish scientists who work on nanotechnology seems to be well connected at the micro level but not so much at the macro level. In other words, researchers tend to collaborate within their own sub-clusters (i.e., groups or universities) more often. The frequencies of the total number of publications that first authors contributed to adhere to Lotka's law:

$$f(y) = .2459 \div y^{1.2881} \quad (1)$$

where $f(y)$ denotes the relative number of authors with y publications (the K-S DMAX = 0.6323) (Rousseau, 1997), indicating that a small number of well-known scientists have stronger positions in the network. As mentioned earlier, although some scientists from smaller universities with the lower degree centrality coefficients have appeared in the network structure as a turning point, one can call them as non-elite authors. However, their impact on knowledge diffusion is remarkable.

⁴ Note that some author names with the same initials are affiliated with two different universities in this period (e.g., Çakmak M at both Koç and Kırıkkale Universities and Turan R at both Ege and Middle East Technical Universities). They may well be the same authors who may have moved from one university to the other during this period.

We also carried out a co-word analysis on the words that appear in the titles of articles extracted from WoS to find out the most frequently used terms between 2000 and 2005, and between 2006 and 2010. The first 75 most frequently occurring words in each period were collected, processed and compiled by the software.⁵ Non-trivial words were eliminated. In order to analyse the word/document occurrence matrix in terms of its latent structure, SPSS software version 16.0 was used to factor analyse the co-occurrence of words. Factor analysis maps each word to a different component (research strand) with the highest factor loading. SPSS created two factors from the list of the co-words. Table 4 and 5 show the output of factors for the periods of 2000-2005 and 2006-2011 along with the loadings of different words in each factor (not all 75 words listed in the tables). According to eigenvalues, the first factor explains 56% of the variance in the entire data set for the period of 2000-2005 while the second one explains the rest of the variance (44%). For the 2006-2011 period, the first factor explains 35% of the variance in the entire data set while the second and third ones explain 33% and 32% of the variance, respectively.

Table 4. Factor analysis of co-words in the titles of nanotechnology papers (2000 and 2005).

Rotated component matrix ^a			
Words	Factor 1	Words	Factor 2
CHEMICAL	.999	PLASMA	.999
QUANTUM	.999	TREATMENT	.999
STEEL	.998	CONDUCTING	.990
HYDROGEN	.997	CERAMIC	.982
COPOLYMER	.992	SOL-GEL	.982
FIELD	.992	LAYER	.945
PROPERTIES	.984	OPTICAL	.945
ELECTRICAL	.973	SURFACE	.945

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 8 iterations.

We then produced a normalized cosine extraction of the words and mapped the network structure of co-word analysis in each period using Kamada & Kawai algorithm embedded in Pajek (Fig. 3). Words that appear in both periods belong mainly to Multidisciplinary Science and Materials Science. Represented fields in both periods are as follows: Surface Materials (“Doped”, “Alloy”, and “Plasma”); Chemistry and its subfields (“Coating”, “Crystal” “Catalyst”, and “Sol-Gel”); and Physics (“Quantum”, “Dot” and “Nanotube”). It appears that Turkish nanoscientists work primarily in Material Sciences, followed by Physics and, to some extent, Biotechnology.

⁵ We used the software available at <http://www.leydesdorff.net/software/fulltext/index.htm> to create a normalized cosine symmetric co-occurrence matrix of labels.

Table 5. Factor analysis of co-words in titles of nanotechnology papers (2006 and 2011).

Rotated component matrix^a

Words	Factor 1	Words	Factor 2	Words	Factor 3
COPOLYMER	.766	STEEL	.673	DOT	.687
COMPLEXES	.697	WELL	.655	MORPHOLOGY	.676
CRYSTAL	.674	AQUEOU	.651	ADSORPTION	.654
THERMAL	.653	ZNO	.642	ENERGY	.644
SPECTROSCOPIC	.650	PARTICLE	.626	PREPARED	.641
CHARACTERISTIC	.643	MATERIAL	.625	QUANTUM	.620
COPOLYMER	.766	TEMPERATURE	.620	ELECTRICAL	.619
METAL	.636	CELL	.618	MODIFIED	.610

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Discussion and Conclusion

In this paper, we assessed the network structure of nanotechnology papers authored by Turkish scientists between 2000 and 2011. We used the social network analysis techniques and studied the network properties from different perspectives. We first identified the top 15 universities for each period (2000-2005 and 2006-2011) on the basis of centrality coefficients. They played pivotal roles in the dissemination of nanotechnology knowledge in Turkey. We then created the co-authorship network of nanotechnology scientists and analyzed the network properties (coefficients of degree, betweenness, closeness centralities and PageRank) of the top 15 authors in each period. We also used the co-word analysis to identify the major nanotechnology research fields in Turkey on the basis of the co-occurrence of words in the titles of papers.

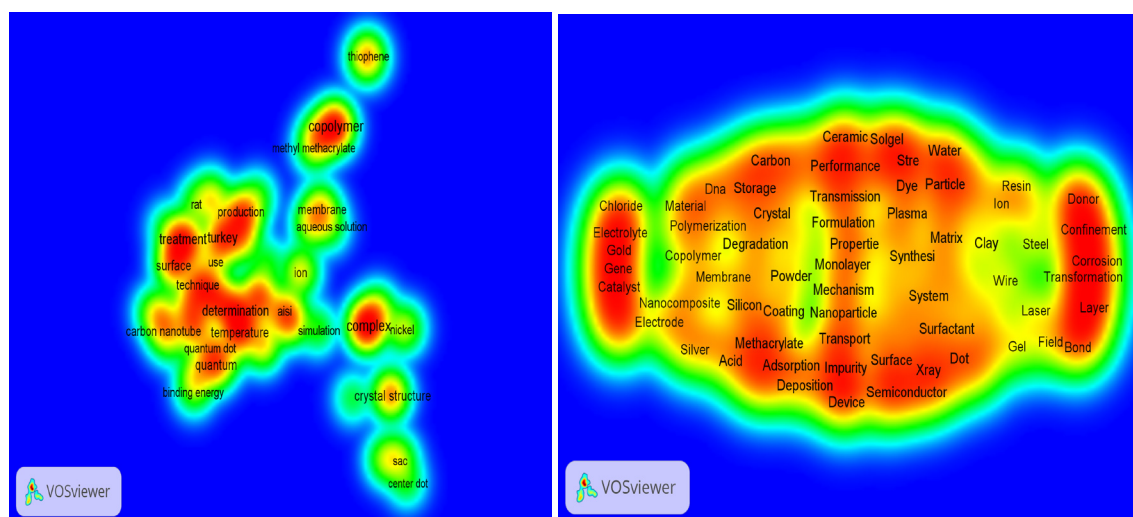


Figure 3. Network of co-word analysis in nanotechnology in Turkey: (l) 2000-2005 and (r) 2006-2011.

Although the number of nodes in the network has increased in the second period (2006-2011), the overall connectedness of the network structures is low. The centrality coefficients of the network structure of the top 15 universities revealed that the social network structure is denser

at the micro level than that at the macro level. While the betweenness centrality remained low and the closeness centrality did not change much, the degree centrality increased almost 60% in the second period, which is an indication of the small world phenomenon in the network structure.

The research output of Turkish nanoscientists and collaboration among them conform to some extent to Lotka's law in that a few researchers tend to publish the bulk of nanotechnology papers while the rest are less prolific. This indicates that Turkish scientists tend to work with prolific authors. The taxonomy identified by the co-word analysis shows that Turkish nanoscientists mainly work in Materials Sciences, Chemistry and Physics. Nanotechnology research continues to flourish due to collaborations at the micro level within the Turkish scientific community and the diffusion of nanotechnology knowledge is accelerating. Bibliometric indicators and network properties reported in this research may help policy-makers to understand the interdisciplinary character of nanoscience and nanotechnology better and develop funding mechanisms accordingly.

References

- Aydogan-Duda, N. (2012). Nanotechnology: A descriptive account. Making it to the forefront in Aydogan-Duda, N. (Ed). *Nanotechnology: A Developing Country Perspective*. 1, (pp. 1-4). New York: Springer.
- Aydogan-Duda, N., & Şener, I. (2010). Entry barriers to the nanotechnology industry in Turkey in Ekekwe, N. (Ed). *Nanotechnology and Microelectronics: Global Diffusion, Economics and Policy*. (pp. 167-173). Hershey, PA: IGI Global.
- Barabasi, AL. & Albert, R. (1999). Emergence of scaling in random networks. *Science Magazine*, 286(5439), 509-512.
- Bozkurt, A. (2015, January). Türkiye, 10 yıldır “en küçük” dünyanın farkında, artık büyük adımlar atması gerekiyor (Turkey is aware of the “smallest” world for 10 years, but it should take big steps). *Bilişim: Aylık Bilişim Kültürü Dergisi*, 43(172), 44-53. Retrieved June 6, 2015 from: http://www.bilisimdergisi.org/s172/pages/s172_web.pdf.
- Burt, R.S. (1992). *Structural Holes: The Social Structure of Competition*, Cambridge, MA: Harvard University Press.
- Callon, M., Courtial, J.P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information*, 22(2), 191-235.
- Centrality. (2015). Retrieved, January 20, 2015, from <http://en.wikipedia.org/wiki/Centrality>.
- Chen, C. (2004). Searching for intellectual turning points: Progressive knowledge domain visualization. *PNAS*, 101(Suppl. 1), 5303-5310.
- Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z., & Pellegrino, D. (2009). Towards an explanatory and computational theory of scientific discovery. *Journal of Informetrics*, 3(3), 191-209.
- Darvish, H. (2014). *Assessing the diffusion of nanotechnology in Turkey: A social network analysis approach*. Unpublished PhD dissertation. Hacettepe University, Ankara.
- Darvish, H. & Tonta, Y. (2015). The diffusion of nanotechnology knowledge in Turkey (submitted).
- Denkbaş, E.B. (2015, January). Nanoteknolojiye yapılacak yatırımlar, ülkelerin ekonomik gücünü yansıtabilecek bir parametre olacak (Investments in nanotechnology will become a parameter reflecting economic powers of countries). *Bilişim: Aylık Bilişim Kültürü Dergisi*, 43(172), 78-87. Retrieved June 6, 2015 from: www.bilisimdergisi.org/pdfindir/s172/pdf/78-87.pdf.
- Freeman, L.C. (2004). *The Development of Social Network Analysis: A Study in the Sociology of Science*. Vancouver: Empirical Press.
- Granovetter, M. (1974). *Getting a Job: A Study of Contacts and Careers*. Cambridge, Mass: Harvard University.
- Kostoff, R.N., Koytcheff, R.G., & Lau, C.G.Y. (2007). Global nanotechnology research literature overview, *Technological Forecasting & Social Change*, 74, 1733-1747.
- Kostoff, R.N., Stump, J. A., Johnson, D., Murday, J.S., Lau, C.G.Y., & Tolls, W.M. (2006). The structure and infrastructure of global nanotechnology literature. *Journal of Nanoparticles Research*, 8, 301-321.
- Martin, T., Ball, B., Karrer, B., Newman M. E. J. (2013). *Coauthorship and citation in scientific publishing*. Retrieved December 27, 2014 from <http://arxiv.org/abs/1304.0473>.
- Moody, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. *American Sociological Review*, 69(2), 213-238.
- Nanotechnology. (2015). Retrieved, January 20, 2015, from <http://en.wikipedia.org/wiki/Centrality>

- Newman, M. E. J. (2000). The structure of scientific collaboration networks. *PNAS*, 98(2), 404–409.
- Otte, E., & Rousseau, R. (2002). Social Network Analysis: A powerful strategy, also for the information sciences. *Journal of information Science*, 28(6), 443–455.
- Özel, B. (2010). *Scientific collaboration networks: Knowledge diffusion and fragmentation in Turkish management academia*, Unpublished PhD dissertation, Bilgi University, Istanbul.
- Özgüz, V. (2013). Nanotechnology research and education in Turkey (presentation slides). Retrieved, December 27, 2014, from: http://rp7.ffg.at/upload/medialibrary/12_Oezguez.pdf
- Page, L., & Brin, S. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30, 107–117.
- Rousseau, R. (1997). *Sitations: an exploratory study*. *Cybermetrics*. Retrieved, February 14, 2014, from <http://cybermetrics.cindoc.csic.es/articles/v4i1p4.pdf>.
- Scott, J. (2000). *Social Network Analysis: A Handbook*. 2nd ed. London: Sage.
- Watts, D. (2003). *Six Degrees: The Science of a Connected Age*. New York: W. W. Norton & Company.
- Watts, D. J. & Strogatz, S. (1998). Collective Dynamics of “small-world” Networks. *Nature*, 393(6684), 440–441.

Analysis of the Spatial Dynamics of Intra- v.s. Inter-Research Collaborations across Countries¹

Lili Wang¹ and Mario Coccia²

¹ wang@merit.unu.edu

UNU-MERIT, Keizer Karelplein 19, 6211 TC, Maastricht (The Netherlands)

² mario.coccia@ircres.cnr.it

CNR -- National Research Council of Italy, 10024 Moncalieri, Torino (Italy)

Abstract

The purpose of this paper is to analyse the evolutionary pattern of international research collaborations. Using publication data from 1997 to 2012, this study decomposes international collaborations into two complementary types, intra-collaboration (within the same geographical area) and inter-collaboration (across different geographical areas). Our results show that the geographical concentration of international research collaborations is reducing. The formation of new network structure of international research collaborations is driven by the increase of inter-research collaborations of countries across different geographical areas rather than intra-collaborations of countries within the same geographical area.

Conference Topic

International collaboration

Introduction

Scientific collaborations have been widely acknowledged to be efficient in managing time and labour in research labs (Coccia, 2014; Solla Price & Beaver, 1966), improving research quality (Presser, 1980; Narin et al., 1991; Katz & Hicks, 1997) and spurring the breakthroughs of scientific research for supporting competitiveness (Coccia, 2012). A number of factors have contributed to the continuous increase of international research collaborations and co-authored papers (Beaver & Rosen, 1978; Frame & Carpenter, 1979; Katz & Martin, 1997). Along with the steady rise of international scientific collaborations, a better understanding on the structure of the global research network across geo-economic areas and its evolutionary pattern are needed for scholars and policy makers.

The high heterogeneity across countries – in terms of size, scientific capacity of the national system of innovation, etc. – generates a variety of patterns of the international research collaborations (Melin, 1999; Narin et al., 1991; Ozcan & Islam, 2014). A main issue in economics of science is to determine how and to which extent countries are engaged in international research collaborations so as to understand the behaviour of knowledge flows and to design research policies for improving the scientific research production which will in turn to enhance national competitiveness.

Luukkonen et al. (1992) maintain that the map of collaborative connections between countries corresponds to a geographical map. Frame et al. (1977, p. 502), considering data of 1973, claim that: “the production of mainstream science is more heavily concentrated in the hands of a few countries”. Hoekman et al. (2010), using data on co-publications in European countries, show that research collaborations are geographically localized and despite a research heterogeneity in European countries in terms of research collaboration patterns, there

¹ Mario Coccia gratefully acknowledges financial support from United Nations University -The Maastricht Economic and Social Research Institute on Innovation and Technology (Contract ID 606U U-04 76) where this joint research was conducted while he was a visiting researcher.

is “a gradual convergence is taking place toward a more integrated interconnected European science system” (Hoekman et al., 2010, p. 672).

The purpose of this research is to investigate the evolutionary pattern of international research collaborations across countries. Emphasis is placed on two complementary collaboration types, i.e. intra- and inter-collaborations. The former refers to research collaborations conducted by countries within the same geographical area; the latter refers to research collaborations engaged by countries from different geographical areas.² Increase of intra-collaborations indicates that cooperation is more and more bounded within certain geographical territories, while increase of inter-collaborations signals the fade of geographical limit.

The main research questions of this paper are:

- How does the distribution of international collaborations across countries evolve over time?
- What type of research collaborations (inter- or intra-) plays a more important role in reshaping the global collaborative scientific network across geo-economic areas?
- How do inter- and intra- connections change in the global collective network?

The analysis of the temporal and spatial evolution of these patterns is of great scientific interest for researchers and policy makers in order to better master knowledge flow and optimize collaborative research output across countries.

Data and methodology

The data of this study are collected from publications in academic journals covered by the Science Citation Index (SCI) and Social Sciences Citation Index (SSCI). In particular, this study refers to dataset by National Science Foundation (2014)-National Center for Science and Engineering Statistics, special tabulations from Thomson Reuters (2013), SCI and SSCI. Collaboration data cover two years 1997 and 2012 and 40 countries (see the list in Appendix A). These 40 countries produce about 97% of the global total articles over 1997-2012. The 40 countries are classified into eight geographical areas: North America, South America, Europe Union, Other Europe, Middle East, Africa, Asia and Australia/Oceania (see Appendix A).

The analysis consists of the following steps:

- Firstly, to analyse the worldwide distribution of international collaborations, this study uses Lorenz curves and Gini coefficient. Lorenz curve is indicated by $L(X)$, then Gini coefficient can be derived as follows:

$$\text{Gini coefficient } (G) = 1 - 2 \int_0^1 L(X) dX \quad (1)$$

G is main indicator of concentration of the distribution of data.

- Secondly, to map the research connections between countries, both absolute collaborative output (number of articles) and collaboration intensity are considered. The former data set demonstrates the major players in the global collaboration research network while the latter puts all countries into one comparable framework. Although the matrix of co-authored papers between countries provides us main information concerning the output co-occurrence, the number of collaborated output might have different meanings for the collaborating country pair due to their different research capacity. For instance, suppose that a research collaborative pair is formed by Country A (of which the number of total publications is 1000) and Country B (of which the number of total publications is 10,000). Collaboration intensity (the ratio of collaborative output to national total publications) presents a stronger collaboration

² The under studied geographical areas are: North America, South America, Europe Union, Other Europe, Middle East, Africa, Asia and Australia/Oceania.

link for country A than B. Therefore, extra caution should be exercised when analysing the collaborative connections between research partners.

Based on eight geographical groups, this study disentangles intra-collaborations (between countries located in the same geographical area) from inter-collaborations (between countries of different geographical areas).³

Salton and Jaccard indexes are both valuable in measuring relative collaboration intensity (cf. Luukkonen et al., 1993). The collaboration index by Salton's measure (CSI) is

$$CSI = \frac{CO_{ij}}{\sqrt{P_i * P_j}} \quad (2)$$

whereas, the Jaccard's measure (CJI) is given by:

$$CJI = \frac{CO_{ij}}{P_i + P_j - CO_{ij}} \quad (3)$$

Where CO_{ij} is the number of co-authored papers between country i and country j

P_i is the total publication number by country i

P_j is the total publication number by country j

In addition, to understand the intra- and inter- collaborations by Salton and Jaccard indices (equations (2) and (3)), the adapted intra- and inter- collaboration intensities are

- $CSI_{intra} = \frac{CO_{ij}}{\sqrt{P_i * P_j}} \quad (i \& j \in \text{same geographical area}) \quad (4)$

- $CSI_{inter} = \frac{CO_{ij}}{\sqrt{P_i * P_j}} \quad (i \& j \in \text{different geographical areas}) \quad (5)$

- $CJI_{intra} = \frac{CO_{ij}}{P_i + P_j - CO_{ij}} \quad (i \& j \in \text{same geographical area}) \quad (6)$

- $CJI_{inter} = \frac{CO_{ij}}{P_i + P_j - CO_{ij}} \quad (i \& j \in \text{different geographical areas}) \quad (7)$

Coefficient of variation is also applied to assess the dispersion of data.

- Thirdly, from a dynamic perspective, this study applies network analysis to explore the structure of international collaborations and its changes from 1997 to 2012. In particular, intra- and inter- scientific ties across countries are distinguished from each other in the networks.

Empirical analysis

Global distribution of scientific research and collaborations

It has been well recognized that research capability and resources are unevenly distributed in the world, and hence scientific research output is concentrated in certain countries which are scientifically strong (Frame et al., 1977). By measuring the statistical dispersion of total publications and international collaborations, Table 1 shows that the Gini coefficient of internationally co-authored papers is lower than that of total publications, which means the former is distributed more evenly across countries than the latter. Most importantly, the Gini coefficients for both types of scientific outputs are decreasing over years. This means that the distributions of total publications and internationally co-authored papers both became less geographically concentrated in the later years.

³ Refer to Appendix A for detailed group information.

Table 1. Gini Coefficient over years

	<i>1997</i>	<i>2002</i>	<i>2007</i>	<i>2012</i>
Total publications	0.67	0.63	0.61	0.59
Internationally co-authored papers	0.60	0.58	0.56	0.54

Dynamics of international collaborations

Salton and Jaccard measures are considered for estimating the collaboration intensity (Figure B1 and B2, in the Appendix B). The arithmetic mean of Salton measure is as twice as that of Jaccard measure, which is in line with Hamers, et al. (1989). However, the coefficient of variation in Jaccard is somewhat higher than that of Salton (see Fig. B1 and B2), indicating a greater dispersion of collaboration intensities is measured by Jaccard index. As the aim of this study is to analyse collaborative research variability between countries, intensities derived from Jaccard index seem to be more suitable.⁴

At the level of geographical groups, Figure 1 shows the relationship of the intra- and inter-collaboration intensities between 1997 and 2012. Red dots represent the inter-collaboration intensity and green ones represent intra-collaboration intensities. A dot being above diagonal line indicates that the collaboration intensity of this observed unit has increased in 2012 in contrast to that of 1997. Likewise, a dot underneath the diagonal indicates that the international collaboration intensity has decreased in 2012 compared to that of 1997. The fact that all the dots lying above the diagonal line suggests that both intra- and inter-collaboration intensities in all geographical areas have improved over years. On the other hand, by comparing the red and green dots, it is of great interest to observe that inter-collaborations in all geographical areas have increased dramatically while intra-collaborations stay mostly low and close to the diagonal line. The intra-collaboration intensity in the European Union (EU) is the only exception with high level of intra-collaborations in both 1997 and 2012, which is a phenomenon of “Europeanisation” as discussed by Mattsson et al. (2008). In general, this figure shows that intra-collaborations tend to be static while inter-collaborations exhibit high dynamics of growth.

⁴ In the rest of the paper, we present only results calculated based on Jaccard measure. Similar results using Salton measure are available upon request.

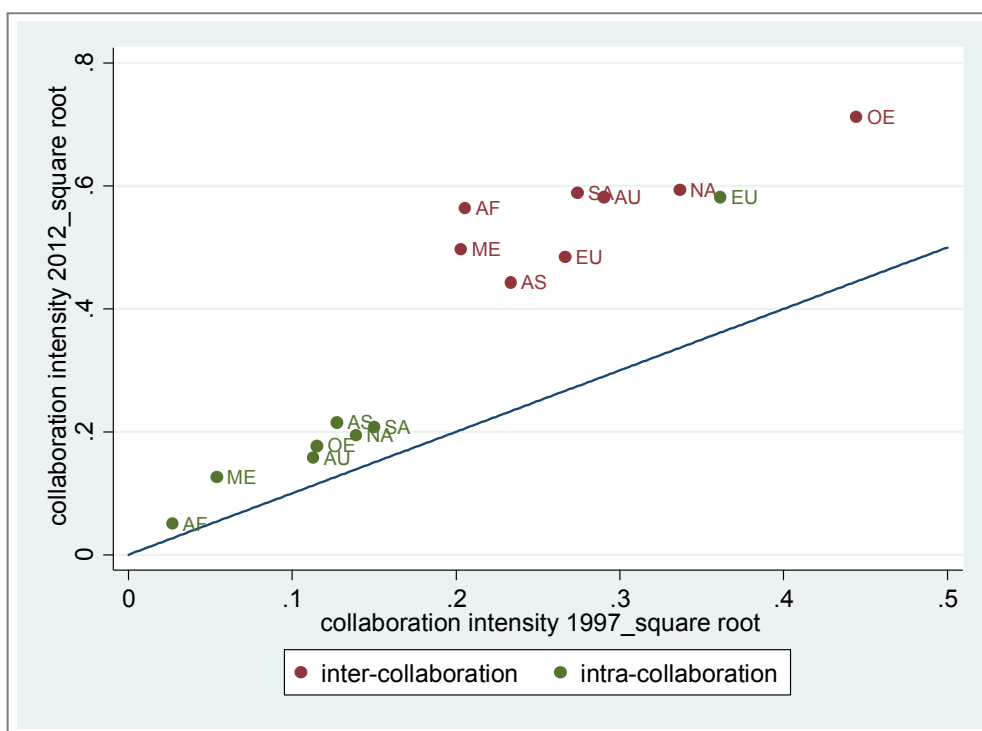


Figure 1. Comparison of international collaboration intensity (inter vs. intra)

Note: 1) The eight geographical areas are: North America (NA), South America (SA), European Union (EU), Other Europe (OE), Middle East (ME), Africa (AF), Asia (AS) and Australia/Oceania (AU). 2) Collaboration intensity is measured by Jaccard index.

To further understand the changes of collaborative performance in individual countries, Figure 2 presents the intra- and inter-collaboration intensity in the 40 under studied countries. Countries in European Union are the only ones showing growth of both intra- and inter-collaborations. This can be the result of European Commission's policy which stimulates cooperation between European countries. In the rest countries, the intra-collaboration performance looks all static, while inter-collaborations have risen obviously. Among all the countries, a group of Asian countries (China, India, Japan, Singapore, and South Korea) show relatively slow growth in inter-collaborations.

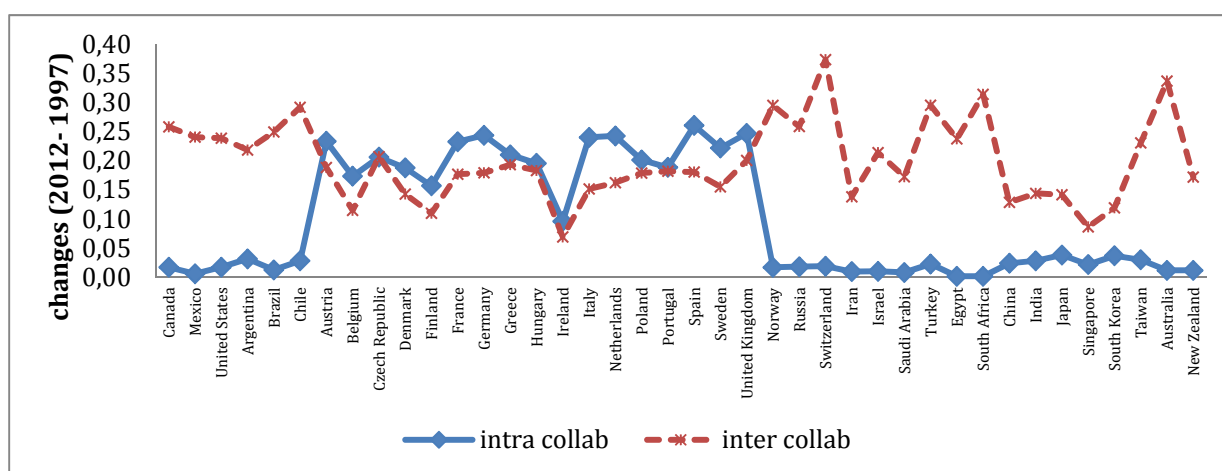


Figure 2. Changes of international collaboration intensity by country (inter vs. intra)

Note: 1) Collaboration intensity is measured by Jaccard index. 2) The value of y-axis is calculated by the collaboration intensity in 2012 minus that in 1997.

Networks of research collaborations

Based on Jaccard collaboration intensity, collaborative networks across 40 countries in 1997 and 2012 are provided in Figure 3 and 4. The thickness of each edge between two nodes reflects the strength of their collaborative relationship. The higher collaboration intensity one country pair has, the thicker their connection line is. In order to distinguish between intra- and inter-collaborations, geographical areas are presented in different colours.⁵ Lines connecting nodes in different colours represent inter-collaborations, while those between nodes in same colours represent intra-collaborations. The size of each node embodies its aggregated collaboration intensity (including both intra- and inter-collaborations).

Figure 3 shows that scientific collaboration networks have been, to some degree, formed by geographic ties. Apart from the intensive connections between European countries (intra-collaborations), there are a few geographically biased small clusters are of great interest. The rectangular cluster in Nordic countries (formed by Denmark, Sweden, Norway and Finland) and the triangular cluster in South America (formed by Chile, Brazil and Argentina) both indicate that scientific collaborations are geographically localized. Besides these small clusters, in North America, a strong tie is observed between United States and Canada. In Asia, China is mainly connected with Japan. In Australia/Oceania, New Zealand has a strong connection only with Australia.

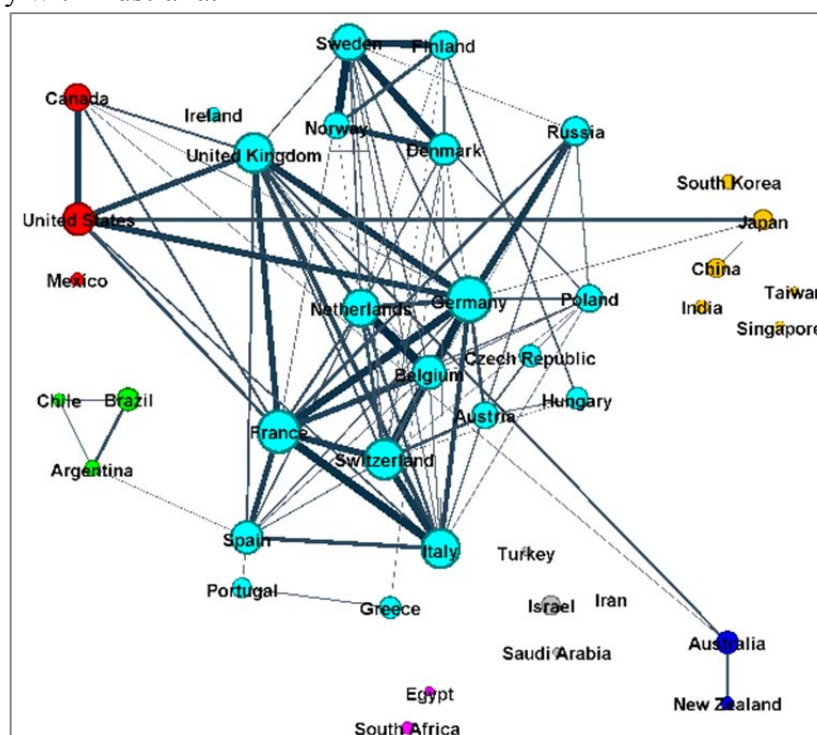


Figure 3. Network of global research connections in 1997.

Note: 1) A filter of 0.0083 is applied in this figure, which means that edges with collaboration intensity less than 0.0083 are omitted. 2) The thickness of each edge between two nodes reflects the strength of their collaborative relationship. 3) The size of each node embodies its aggregated collaboration intensity.

⁵ To emphasize the effect of geographical locations, *European Union* and *Other Europe* are regarded as one group in the network figures (Fig. 3 and 4).

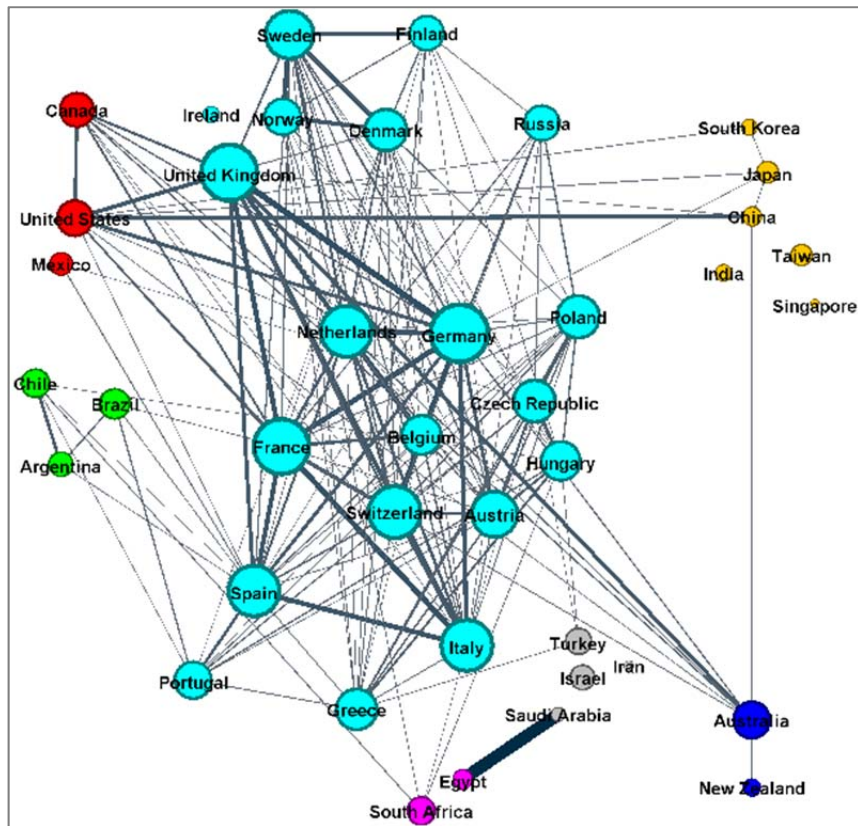


Figure 4. Network of global research connections in 2012.

Note: 1) The network in 2012 is much denser than that of 1997. In order to keep the visualization compact and readable, filter applied in this figure is as twice high as the 1997 figure. Edges with collaboration intensity less than 0.016 are omitted. 2) The thickness of each edge between two nodes reflects the strength of their collaborative relationship. 3) The size of each node embodies its aggregated collaboration intensity.

In order to understand the dynamics of international collaborations, it is necessary to compare the structure of networks in the earlier year 1997 (Fig. 3) with that of the later year 2012 (Fig. 4). In contrast with 1997, the aggregated collaboration intensity (embodied by the circle size of each node) for most countries has increased in 2012. In particular, an important observation is that, the variety of inter-collaborations (lines between different coloured nodes) has grown significantly in 2012, while the connection strength between major intra-collaborative partners (nodes with the same colours) stayed roughly at original level of 1997.

In contrast with the structure in 1997 (Fig. 3), the rectangular Nordic cluster and triangular South American cluster in 2012 have both increased their inter-connections with countries beyond their geographic neighbours (see Fig. 4). The strong tie between Chile and Brazil (i.e. intra-collaboration) has been weakened while both Chile and Brazil developed new inter-collaborative partnerships with countries from other geographical areas. Similarly, the tie between Finland and Denmark became relatively weaker, whereas both of them established more connections with various countries. Due to the effect of “Europeanisation” of this geo-economic area, the new major collaboration partners are still within Europe, but far beyond the old Nordic limit in the later year.

Asian countries, though still with relatively low collaboration intensity, have increased scientific cooperation with the United States (i.e. known as type of inter-collaborations). In particular, China has developed a very strong collaborative tie with the United States and a reasonable partnership with Australia, which are both inter-collaborations. Yet as the second

largest producer of scientific publications, China did not develop any new strong collaborative ties (i.e. intra-collaborations) within its own geographical area.

Located in North America, Mexico seemed to have developed new collaborative research partners only beyond its own geographical area (i.e. inter-collaborations). As one of the most dynamic countries regarding international research collaborations, South Africa seemed to have built inter-collaborative relationships mainly in Europe and South America. Different from the isolated situation in the earlier stage (1997), Egypt and Saudi Arabia developed an extremely strong research partnership in 2012.⁶ Their connection with each other was so strong that they hardly had any cooperation with any third countries.

Conclusions

The main lessons learned of this research can be synthesized as follows:

- 1) The Gini coefficients for total publications and collaborations were both smaller in 2012 than 1997, indicating that the distribution among the under studied 40 countries became more and more balanced. Nevertheless, it is worthwhile to note that the distribution of total publications was more divergent than that of internationally co-authored papers.
- 2) In the process of evolution of international collaborations, evidence shows significant difference between intra- and inter- collaborations. In all geographical areas, except European Union, the intra collaboration performances exhibited a steady-state pattern, whereas inter-collaborations in the global network research structure have risen dramatically.
- 3) From a dynamic point of view, the comparison of 1997 and 2012 networks shows that inter-collaborations (between countries from different geographical areas) have grown significantly in the later stage, while the connection strength between major intra-collaborative partners stayed mostly unchanged. This finding indicates that recent research network across countries has a higher global inter-connection beyond geographical territories, which is likely driven by advances of ICT and transportation new technologies and improvement of socio-economic systems.

In short, the increase of research collaborations between countries from different geographical areas has reshaped the global structure of international scientific collaborations. In the modern process of knowledge production, countries seem to be looking for more diverse collaborative partners worldwide.

References

- Beaver, de B.D. & Rosen, R. (1978). Studies in scientific collaboration – Part. I. The professional origins of scientific co-authorship, *Scientometrics*, 1, 65-84.
- Coccia M. (2010). Democratization is the Driving Force for Technological and Economic Change, *Technological Forecasting & Social Change*, 77, 248-264.
- Coccia, M. (2012). Political economy of R&D to support the modern competitiveness of nations and determinants of economic optimization and inertia, *Technovation*, 32, 370-379
- Coccia M. (2014). Driving forces of technological change: The relation between population growth and technological innovation-Analysis of the optimal interaction across countries, *Technological Forecasting & Social Change*, 82, 52-65
- de Solla Price D. & de B. Beaver D. (1966). Collaboration in an invisible college, *American Psychologist*, 21, 1011-1018.
- Frame J. D. & Carpenter M. P. (1979). International research collaboration, *Social Studies of Science*, 9, 481-497.
- Frame, J.D., Narin, F. & Carpenter, M.P. (1977). The distribution of world science, *Social Studies of Science*, 7, 501-516.

⁶ Although Egypt and Saudi Arabia are classified into different groups, they are in geographically adjacent. Therefore their collaborative relationship can be still regarded as a result of geographical localization.

- Hamers, L., Hemeryck, Y., Herweyers, G., Janssen, M., Keters, H., Rousseau, R., & Vanhoutte, A. (1989). Similarity measures in scientometric research: the Jaccard index versus Salton's cosine formula. *Information Processing & Management*, 25(3), 315-318.
- Hoekman, J., Frenken, K. & Tijssen, R. (2010). Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy*, 39, 662-673.
- Katz, J. S. & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26, 1-18.
- Katz, J.S. & Hicks, D. (1997). How much is a collaboration worth? A calibrated bibliometric model. *Scientometrics*, 40, 541-554.
- Luukkonen, T., Persson, O., & Sivertsen, G. (1992). Understanding patterns of international scientific collaboration, *Science, Technology & Human Values*, 17, 101-126.
- Luukkonen, T., Tijssen, R.J.W., Persson, O., & Sivertsen, G. (1993). The measurement of international scientific collaboration. *Scientometrics*, 28, 15-36.
- Mattsson, P., Laget, P., Nilsson, A. & Sundberg, C-J. (2008). Intra-EU vs. extra-EU scientific co-publication patterns in EU. *Scientometrics*, 75, 555-574.
- Melin, G. (1999). Impact of national size on research collaboration. *Scientometrics*, 46, 161-170.
- Narin, F., Stevens, K. & Whitlow, E.S. (1991). Scientific co-operation in Europe and the citation of multinationally authored papers, *Scientometrics*, 21, 313-323.
- Ozcan, S. & Islam, N. (2014). Collaborative networks and technology clusters –The case of nanowire. *Technological Forecasting and Social Change*, 82, 115-31.
- Presser, S. (1980). Collaboration and the quality of research. *Social Studies of Science*, 10, 95-101.

Appendix A. Country/economy of the sample

nr	country	Geo-Economic Area
1	Canada	North America
2	Mexico	
3	United States	
4	Argentina	South America
5	Brazil	
6	Chile	
7	Austria	European Union
8	Belgium	
9	Czech Republic	
10	Denmark	
11	Finland	
12	France	
13	Germany	
14	Greece	
15	Hungary	
16	Ireland	
17	Italy	
18	Netherlands	
19	Poland	
20	Portugal	
21	Spain	
22	Sweden	
23	United Kingdom	
24	Norway	Other Europe
25	Russia	
26	Switzerland	
27	Iran	Middle East
28	Israel	
29	Saudi Arabia	
30	Turkey	
31	Egypt	Africa
32	South Africa	
33	China	Asia
34	India	
35	Japan	
36	Singapore	
37	South Korea	
38	Taiwan	
39	Australia	Australia/Oceania
40	New Zealand	

Appendix B:

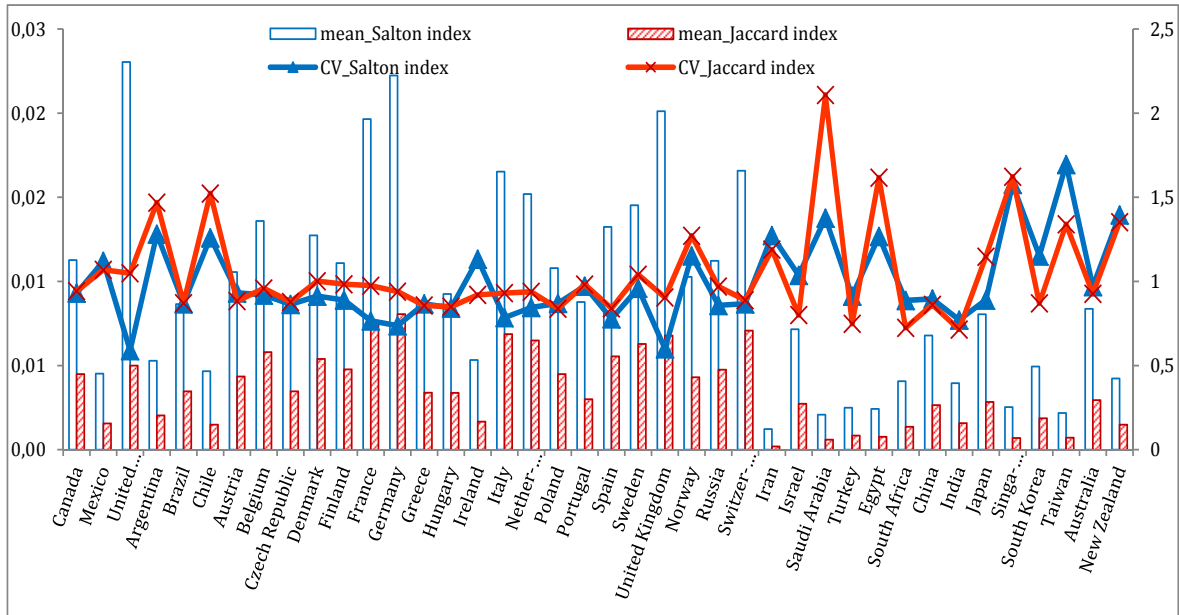


Figure B1. Mean and coefficient variation for collaboration indices (Salton vs. Jaccard) 1997

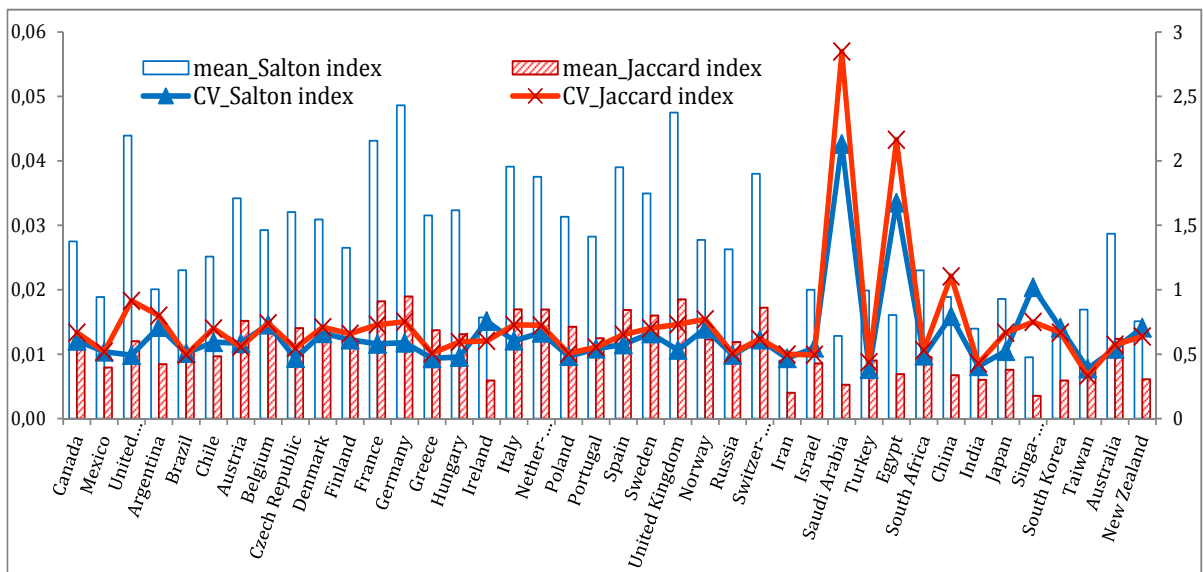


Figure B2. Mean and coefficient variation for collaboration indices (Salton vs. Jaccard) 2012

Nanotechnology Research in Post-Soviet Russia: Science System Path-Dependencies and their Influences

Maria Karaulova¹, Oliver Shackleton¹, Abdullah Gök¹ and Philip Shapira^{1,2}

¹ Manchester Institute of Innovation Research, Manchester Business School, University of Manchester (United Kingdom)

² School of Public Policy, Georgia Institute of Technology (USA)

Abstract

This paper contributes to the analysis of Russian research dynamics and output in nanotechnology. The paper presents an analysis of Russian nanotechnology research outputs during the period of 1990-2012. By examining general outputs, publication paths and collaboration patterns, the paper identifies a series of quantified factors that help to explain Russia's limited success in leveraging its ambitious national nanotechnology initiative. Attention is given to path-dependent institutionalised practices, such as established publication pathways that are dominated by the Academy of Sciences, the high centralisation of the entire research system, and issues of internal collaborations of actors within the domestic research system.

Conference Topic

Country-level Studies

Introduction

Nanotechnology has been an interest of bibliometric research since the early 2000s after the United States and China adopted large-scale policy and funding programmes to stimulate scientific development by massively investing in this interdisciplinary research area. China has been among the countries with a large increase in research outputs in nanotechnology, and is the emerging economy that is frequently the focus of researchers (Appelbaum et al., 2011; Bhattacharya & Bhati, 2011; Liu et al., 2009).

Other emerging and transitional economies have also invested in nanotechnology development. Russia is a particular case among these countries, because the National Nanotechnology Initiative that was adopted in 2007 was a political as well an economic, scientific and technological project. The Russian government picked up on global trends and invested greatly in development of nanotechnology. On a purchasing power basis, it is suggested that public investment in Russian nanotechnology has rivalled that of the US and China (Schiermeier, 2007). Lux Research (2013) estimates that Russian nanotechnology investment has consistently been the third largest in the world after the US and China: Russia invested over \$1 bln in 2010 and 2011 in nanotechnology projects, and just under \$1 bln in 2012. However, with lower than anticipated results in nanotechnology, the Russian government has decreased its investment programme and the share of Russia in world nanotechnology funding dropped from 15% to 13% in 2013. It is anticipated to continue decreasing.

Important changes and structural reforms of Russian science (including nanoscience) have been implemented only relatively recently, in the mid- to late-2000s, almost two decades after the dissolution of the Soviet Union in 1991. Until then, Russian science was relatively unchanged from rules and institutional developed during the Soviet era. The Academy of Sciences of Russia maintained its Soviet-style organisation up until 2013 when it was subjected to a radical reform. Universities were reformed in 2008 and 2009 to move them away from mainly teaching and to develop research capabilities and to try to emulate US research clusters. The funding structure for Russian science was tied to four-year umbrella research programmes accompanied by small-scale research foundations until 2013, when decisions were made to reform Russia's Federal Targeted Programmes and Grant

Programmes towards more grant-based system. Importantly, the Russian National Nanotechnology Initiative and the associated surges in interest and investment pioneered the system-wide initiatives that started several years before other large-scale top-down changes.

Existing literature on nanotechnology research and innovation in Russia is less prodigious than for other “Rising Powers” countries, particularly China but also including Brazil and India. Scientometric analyses often examine Russian nanotechnology development as a benchmark for other emerging economies, mainly China and India (Liu et al., 2011, 2009) rather than deeply probing within the Russian system. At the same time, there is an important strand of scientometric work on Russian science and technology (including nanotechnology) produced by the Russian research community itself. In these cases, research is often descriptive or addresses internal debates within Russia (Terekhov, 2012, 2011), and sometimes lacks a critical approach. Additionally, most of these studies remain mostly background reference country reports (and are frequently only available in Russian).

There are, of course, some exceptions. For example, Klochikhin (2012) contextualised Russian nanotechnology policy in terms of post-Soviet path-dependencies and asked whether it was possible to break out from technological inertia to a new development trajectory. There are other studies of Russian nanotechnology that pose similar questions, be it from the industry and market formation perspective (Ananyan, 2005), or regulation (Gokhberg et al., 2012). A recent overview of the Russian Science, Technology and Innovation system (Karaulova et al., 2014) provides background for discussion of persisting path-dependencies. In the present paper, we build on, and extend, this prior work to examine Russia’s technology development policies and to reflect on the challenges posed by its persistent and deeply-embedded path-dependent practices.

Data and Methodology

The dataset for our research covers the time period from 1990 to 2012, which includes the transitional period after the breakup of the Soviet Union, the Russian Nanotechnology Initiative (NNI) development (2004 – 2007) and the post-NNI period of nanotechnology research. We first provide an updated profile of nanotechnology research in Russia since the breakup of the Soviet Union until 2012. Second, we investigate the possible emergence of new trends of research of Russian nanotechnology after the adoption of large-scale policy programs. Third, we use self-reported publication data in order to illustrate the path-dependent nature of Russian nanotechnology research.

The bibliometric analysis draws on datasets of nanotechnology publications and patents developed by researchers at Georgia Institute of Technology and the Manchester Institute of Innovation Research. Two data sources are used: the Web of Science (scientific publications) (WoS) and Derwent Innovations (patents). Both data sources are published and made available in the Web of Knowledge by Thomson Reuters. Nanotechnology records in the databases are identified using the two-stage search strategy detailed in Porter et al. (2008), and updated in Arora et al. (2012). A keyword search based on a Boolean query is applied. Unrelated records are then removed by applying exclusion terms.

The defining characteristic that we used to identify Russian publications was that at least one author of each included publication had to have a Russian affiliation address (Soviet Union in 1990-1992; Russia subsequently). The primary language of publications in the dataset is English, but specialised editions that include translated articles originally published in Russian are included as well. In total 33,538 Russian nanotechnology publication records were identified in 1990-2012. We acknowledge that there are limitations in using WoS for capturing the totality of Russian science activity (but see also subsequent discussion in this paper of Russian journal publishing strategies).

A feature of the Soviet Union, carried over into the Russian Federation, is that science was and is developed in parallel – but not always in cooperation – with researchers elsewhere in the world. This influences the choice of terminology used by Russian researchers. For example, it has been observed that there is a rich tradition of nanotechnology research in Russia. Alexander Terekhov traces the technological development of Russian nanotechnology back to 1980s when the understanding of the physical properties of ultra-dispersed states enabled Soviet researchers to construct the first lasers and to conduct experiments at the nano-scale (Terekhov, 2013). But the term nanotechnology was not necessarily used at that time. A simple search strategy would not pick up on many Russian nanotechnology publications, especially in earlier years, which are crucial to understand trends of overall growth and development. We judge that the more complex and nuanced approach we apply is better able to capture the emergence and development of the Russia nanotechnology field.

After the publication data was collected and cleaned from unrelated records, further data cleaning to remove duplicates and consolidate organizational and author names was undertaken using VantagePoint text mining software. Cleaning is a large part of our methodology. One of the biggest problems of country report studies that use bibliometric analysis is the issue of varied affiliation reporting. We have addressed various problems through intense cleaning of the data. One problem of aggregation relates to affiliation (location, funding source, author) categories that the database recognizes as separate, but are actually the same. This is an issue that occurs in the self-reported semi-structured publication data. There are variations in reporting of affiliation data, different ways to spell the name of the organization, abbreviations and others. If left unchallenged, the data may be potentially distorted: the contributions of certain actors may appear as less than it reality, which can be misleading. Another major cleaning issue is disambiguating terms that were lumped together. For example, the process of disambiguation of the “Tech Univ” field and further aggregation of the items highlighted that the original very general field contained mainly records published in three large technical universities, and in a number of smaller ones. Table 1 illustrates examples of the data cleaning strategy.

Table 1. Affiliation Cleaning Strategy Examples.

	<i>Original Record</i>	<i>Cleaned Record</i>
Reporting Style	1. RAS, AM Prokhorov Gen Phys Inst; 2. Russian Acad Sci IOF RAN, Prokhorov Gen Phys Inst;	RAS Inst Gen Phys Prokhorov
Abbreviation	1. MISIS 2. State Univ Moscow Inst Steel & Alloys	Natl Univ Sci & Technol MISIS
Spelling	1. Alfa Akonis Res & Devices Enterprise 2. Alpha Akonis R&D Enterprise	Alpha Akonis R&D Enterprise
Change of Name	1. Leningrad State Tech Univ 2. St Petersburg State Tech Univ	St Petersburg State Tech Univ
Disambiguation	Tech Univ	1. St Petersburg Tech Univ 2. Tech Univ Moscow Inst Elect Technol 3. Tech Univ Berlin

Excessive aggregation of the data may lead to the loss of informative value. The Russian Academy of Sciences (RAS) presents the greatest challenge here. RAS is a large research

organisation that possesses more than 500 research institutes. However, the reported RAS affiliations are disordered, because research institutes often have long names and some of them do not issue guidelines for official English versions. Aggregating all these institutes under the domain of the “Russian Academy of Sciences” would yield analytical benefits in some circumstances, such as broad benchmarking. However, such a large agglomeration is not useful for detailed analysis. In our analyses of nanotechnology publications associated with RAS, we undertook disambiguation and identified 263 distinct affiliations, including research institutes of RAS, scientific centres and observatories.

We further grouped the data according to country, region, and type of affiliation. *Academy of Science* organisations are specific research entities that have wide government affiliations and heavily rely on government funding, that have a wide regional structure and hierarchical administrative division. We separately distinguished *Universities*. *Public Research Organisations* are private and state-owned research institutes that are neither academy of science institutions, nor universities. These also include research foundations and ministries. *Corporate actors* are privately and state-owned company affiliations. Organisations were usually labelled as ‘corporate’ actors if they had a distinctive property type word in their names (LLC, Ltd, GmbH, ZAO etc). *Other* included all other organisations that could not be attributed to any other category

In order to examine the internationalisation of Russian science we also separated publications into nationally collaborated publications (NCP) and internationally collaborated publications (ICP). The two groups are mutually exclusive and highlight the degree to which research produced in Russia only involves domestic actors (NCP), or there are also international partners (ICP).

Table 2. Grouping Results, number of publications.

<i>Internalisation</i>			<i>Domestic Affiliation Groups</i>			
				Orgs	Pubs	Share
			Acad of Sciences	3+1(263)	22927	68.5%
NCP	19098	56.9%	University	396	13868	41.4%
ICP	14440	42.8%	PROs	432	3781	11.3%
			Corporate	420	982	2.9%
			Other	3	3	0%

Results

The annual output of Russian nanotechnology publications steadily increased between 1990 and 2012. In 1998, there was a considerable jump in the number of publications; this probably reflects the fresh inclusion of a series of Russian journals within the WoS. Growth rates for domestic and international publications are almost identical starting from 1999 until 2012 and are about 1.1% per year. On average, domestic publications grow 2% faster than internationally collaborated publications.

The Academy of Sciences, 15 universities and four State Research Institutes are the leading organisations in terms of publication output. Some 68% of domestic publications are produced by the Russian Academy of Sciences and another 12% by Moscow State University. The top five organisations produced together 80% of all publications in 1990-2012 (Table 3). The top three organisations (RAS, MSU and St Petersburg State University) produced 78% of all publications. RAS is the dominant actor in producing nanoscience publications. However, in terms of annual publication outputs, university researchers have been catching up with RAS in the past decade.

Table 3. Biggest Publishers in Russian Nanoscience, 1990-2012.

	<i>Organisation name</i>	<i>Publications</i>	<i>Share</i>
1	Russian Academy of Sciences	22794	68.12%
2	Moscow MV Lomonosov State University	4007	11.98%
3	St Petersburg State University	1208	3.61%
4	Russian Research Centre Kurchatov Institute	613	1.83%
5	Nizhnii Novgorod State University	496	1.48%

Disambiguated, the bibliometric map of Russian science demonstrates a more nuanced picture of interactions in the nanotechnology research (Figure 1). One major research organisation, RAS Institute of Physics and Technology n.a. Ioffe, is a focal point for connecting various regional groupings of research centres, such as a cluster of four RAS institutes on Siberia that closely collaborate with one another, but do not have strong external links.

In terms of research performance, nanotechnology publications that only have Russian authors are cited on average 2.5 times per publication. Out of all domestic actors Russian Academy of Sciences publications collect the highest number of citations: 4.55 p/p. PRO publications, albeit being much smaller in number, collect 3.86 citations p/p. Universities collect on average 3.24 citations p/p, and publications produced by corporate actors collect 2.44 citations p/p.

Table 4. Shares of ICP and Average Citation Rate of Russia's Main Collaboration Partner Countries, 1990-2012.

<i>Country</i>	<i>Germany</i>	<i>USA</i>	<i>France</i>	<i>UK</i>	<i>Japan</i>	<i>Sweden</i>	<i>Italy</i>
ICP %	12.3%	8.2%	5.04%	3.4%	2.9%	2.08%	1.9%
Avg Cit	7.7	9.2	5.8	12.2	6.9	6.04	5.3
	<i>Ukraine</i>	<i>Poland</i>	<i>Spain</i>	<i>Netherlands</i>	<i>Belarus</i>	<i>Finland</i>	<i>South Korea</i>
ICP %	1.8%	1.5%	1.5%	1.4%	1.3%	1.1%	0.9%
Avg Cit	2.4	3.9	5.1	18.9	3.8	4.05	3.9

Patterns of international collaboration seem to be connected to these structural differences. The average number of internationally cited publications is 4.33 times: international collaboration increases average citation by 1.7. There are, however, some regional variations in international collaboration performance outputs (Table 4). Russian international collaborations have strong European orientation, and there is evidence of recurrent path-dependent practices. It is noticeable that former Soviet states and influenced territories, such as Ukraine, Poland and Belarus factor highly in collaborative research. It implies research links are built on the older networks than the current political system and research takes place through these interactions. An impeding factor may be than average citation rates for these countries are significantly lower than for other countries with the same collaboration intensity (refer to Table 4). These 8.3% of CIS-collaborated ICPs represent collaboration patterns that may be detrimental to Russian science.

In the next section we pay particular attention to three elements of nanotechnology research that can highlight path-dependent dynamics of scientific knowledge production in Russia. We define them as journal gatekeepers, centralisation, and institutional diffusion. These all relate to structural features of the Russian science system that have persisted even after the Soviet Union broke apart.

Journal Gatekeepers

The data for journals in which Russian co-authored publications can be found, is available for 32844 publications, which constitutes 97% of the data. The majority of Russian publications in English were published in translated journals. Out of the top-10 journals with the biggest number of Russian publications, 7 are translated versions of Russian journals (refer to Table 5).

Translated versions of Russian journals are identified not by the publishing body (the rights to publish in most cases are owned by Springer), but by the contents of the journal and the editorial board. For example, Springer publishes *The Physics of the Solid State*. The description on the website says “The journal Physics of the Solid State presents the latest results from Russia’s leading researchers in condensed matter physics at the Russian Academy of Sciences and other prestigious institutions” (Springer, n.d.). An analogous journal, called *Phyzika Tvyordogo Tela (The Physics of the Solid State)* is published in Russian by the Ioffe Institute in St.Petersburg (Ioffe Physical Technical Institute, n.d.). The Chief Editor of both journals is A.A. Kaplyanskii, and the editorial board matches both journal records. Tables of contents of issues match as well. Based on these we drew a conclusion that *The Physics of the Solid State* is a translated version of *Phyzika Tvyordogo Tela*, and the ‘publishing body’ is therefore an Institute within the Russian Academy of Sciences (the publishing body of the original), not Springer (the publishing body of the translated version). By doing manual analysis of the top journals in which Russian scientists publish we have identified that at least 25% of the entire publication volume was published in this manner (input of the Russian translated journals in the top-20 journal contributions). The overall contribution of the top-20 journals was 25%.

A paper is first published in a Russian peer-reviewed journal, and subsequently translated and published in the English version without an additional peer review. But it would also depend on the domestic peer reviewer whether a submitted article would be considered for publication and further translation for a WoS-indexed version of a journal. The publisher and the editorial board become important. As Table 5 demonstrates, vast majority of the translated Russian journals are published by the Russian Academy of Sciences and editorial boards mainly consist of members of RAS. This *status quo* is grounded in history: many of them were founded during the Soviet Union to inform the world about achievements of Soviet science.

Table 5. Top 20 Journals of Russian Nanotechnology.

	<i>Journal</i>	<i>Publishing Body</i>	<i>Records</i>	<i>Share</i>
1	Physical Review B	APS	1595	4.86%
2	Physics of the Solid State	RAS	1412	4.30%
3	Semiconductors	RAS	1255	3.82%
4	Technical Physics Letters	RAS	848	2.58%
5	JETP Letters	RAS	828	2.52%
6	Inorganic Materials	RAS	511	1.56%
7	Applied Physics Letters	American Institute of Physics	510	1.55%
8	Journal of Applied Physics	AIP Publishing	505	1.54%
9	Journal of Experimental & Theoretical Physics	RAS	490	1.49%
10	Russian Chemical Bulletin	RAS	411	1.25%

After the breakup of the Soviet Union, these established publication pathways and journals have been maintained and there has not been much impetus for change. Although an opportunity opened for Russian researchers to submit research publications to leading international journals, existing publication practices have persisted. Moreover, temporal dynamics highlight an increasing gap between publications submitted to translated Russian journals and international journals: the difference rose from twice as many translated journal publications as international journal publications in 2000 to 2.67 times in 2005 and to 3.8 times in 2011. In the earlier period this could have been explained by the lack of experience of researchers to publish abroad, or by poor knowledge of English. In the later period the English language problem continues, but it also has become prominent that internal domestic recognition for a Russian researcher can be even more important than international recognition in order to develop and continue a research career in Russia. Therefore, publishing in top domestic journals becomes a priority, and the English translation of these papers in journals that collect few citations is a by-product rather than the goal, because this research is anchored in Russian scientific discourse and debates.

RAS maintains the monopoly over acceptance of research outputs to the leading domestic journals, thus acting as a quality control body. It is also a gatekeeper in the Russian research system as to which domestic researchers are highlighted for international recognition. The domination of the Academy of Sciences constrains other research performers, such as universities and PROs, to develop and take advantage of publicly-provided research resources, for example through the Russian NNI. As a comparison, in their study of Chinese publication patterns Zhou and Leydesdorff (2006) recognised this ‘gatekeeping’ role as one of the main barriers to internationalisation of Chinese science in the early 2000s. However, this pattern has now changed with the emphasis in China in publishing directly in WoS journals.

Centralisation and the Academy

In our analysis, we observe two centralisation trends in publications within the Russian Academy of Sciences. The first of these is *geographical centralisation*. RAS has institutes in all 83 regions of Russia, but four regions (Moscow, St Petersburg, Novosibirsk, and the Moscow Region) produced the largest shares of publications in 1990-2012, contributing over 80% of the total amount. Moscow is the leader with almost 35% of all publications, together with the Moscow Region the agglomeration produced 45.2% of all Academy of Sciences publications. Previously, the high concentration of research in a limited geographical area and

with a large network of ineffective and low-performing institutes has been suggested to be one of the main reasons for the persistent problems of RAS (Graham, 1998).

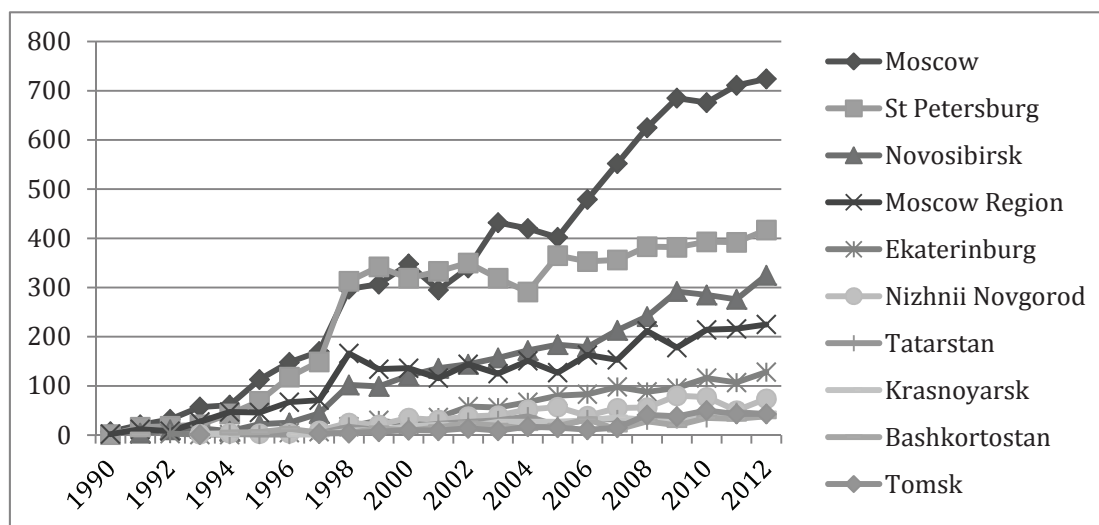


Figure 2. Temporal Dynamics of Geography of Nanoscience in Russia, 1990-2012.

Yet, while problems of RAS centralisation have long been observed, it seems that these trends have intensified in recent years: Academy research is becoming even more centralised (Figure 2). In nanotechnology, RAS institutes in Moscow surged upwards in the mid-2000s, producing almost twice as many publications in 2012 as the research cluster in St Petersburg. Many of these institutes have benefited from recent government science and innovation funding programmes, including specific nanoscience and nanotechnology funding programmes.

The *centralisation of high quality research* is a second persistent trend in Russian nanoscience. RAS has consistently contributed about 70% of the Russian annual publication output. In order to investigate whether quantity translates into quality, we assessed the performance of Russian domestic research system according to the criteria of (1) what affiliations of 10 top-cited (“star”) scientists are, and (2) what affiliations of 100 top-cited publications are.

The top 10 most productive researchers coincide with the most cited researchers, with slight reversal in rank.¹ The majority of these “star” scientists are affiliated with RAS Ioffe Physical Technical Institute in St. Petersburg (Table 6). The Institute itself contributed about 14% of all publications and has an average citation of 6.13. The peak publication activity of all of the most productive scientists was between 1998-2000 after which the decline started. The most productive periods of the most productive Russian nanoscientists coincide with the most productive periods of Russian nanoscience: the contribution of “star” scientists was above 9% in 1996-2001, reaching a peak of 11.5% in 1998. A second, smaller, peak is reached in 2006, after which further decline occurs.

¹ The most highly cited Russian scientists are the ones who collaborated with colleagues at the University of Manchester in a paper in *Science* (Novoselov et al., 2005) that contributed to the award of the 2010 Nobel Prize in Physics to two Manchester researchers. This publication has 3541 citations. To include this exceptionally highly cited publication into the data would overshadow the underlying pattern of Russian nanotechnology performance, so this publication is not included in this part of the citation analysis.

Table 6. "Star" Scientists of Russian Nanoscience.

<i>Rank</i>	<i>Author Name</i>	<i>Affiliations</i>	<i>Times Cited</i>
1	Ledentsov, N	RAS Ioffe Physical Technical Institute	6033
2	Ustinov, Vr	RAS Ioffe Physical Technical Institute	5559
3	Alferov, Zh	RAS Ioffe Physical Technical Institute	5108
4	Kop'ev, P	RAS Ioffe Physical Technical Institute	5052
5	Zhukov, A	RAS Ioffe Physical Technical Institute	3504
6	Valiev, R	RAS Institute of Metals Superplasticity Problems; State Tech Univ of Aviation	3428
7	Egorov, A	RAS Ioffe Physical Technical Institute	2788
8	Morozov, S	RAS Institute of Microelectronics Technology & High Purity Materials	2323
9	Maximov, M	RAS Ioffe Physical Technical Institute	1909
10	Ruvimov, S	RAS Ioffe Physical Technical Institute	1812

The Post-Soviet period saw the rise and the peak of careers of scientists trained in the latter years of the Soviet Union. A drop in productivity coincides with the completion of the active research phase of their careers. There are few new 'rising stars' in the system, which explains the overall decline in performance. This data reinforces concerns about the 'generation gap' in nanotechnology where the average age of researchers is now in the mid-50s (Terekhov, 2011). RAS co-authored 81 out of the 100 most highly cited publications in Russian nanoscience.

Overall, it is notable that RAS dominates in quality as well as the quantity of research in Russian nanoscience. The productivity of RAS reached its peak in the late 1990s and has since then been in decline. The Russian government's support of the development of research universities and RAS reform in 2013 are expected to further contribute to decentralisation of the national research system and to the emergence of new centres of excellence. The trend towards concentration of research in the two capitals – Moscow and St Petersburg – is also a concern as government support to develop scientific research in other regions is limited.

Institutional Diffusion

The third and the final collaboration trend reflects the institutional diffusion of the Russian research system. Institutional theory proponents argue that institutions last and prosper when other elements of the system are dependent on them, e.g. when institutions are diffused well with other institutions (Clemens & Cook, 1999). In a research system this mainly takes form of inter-institutional collaborations. In order to examine the institutional relationships of the Russian research system we investigated (1) whether each organisation preferred to publish on its own; (2) if research was done through the collaboration of authors in one organisation; (3) whether the organisation engaged in collaborative activities with other organisations of the same type; (4) if organisations collaborated nationally; and (5) whether organisations collaborated internationally.

The results of this analysis demonstrate various patterns of domestic collaboration (Figure 3). For instance, corporate publishers have to rely heavily on collaborations, so they have higher rate of collaborations with all types of actors than the average. An asymmetric relationship among the system actors reflects institutional domination of the Academy of Sciences of Russia. The analysis of institutional collaboration patterns demonstrates that there are very weak collaboration links between the Academy of Sciences and other system actors.

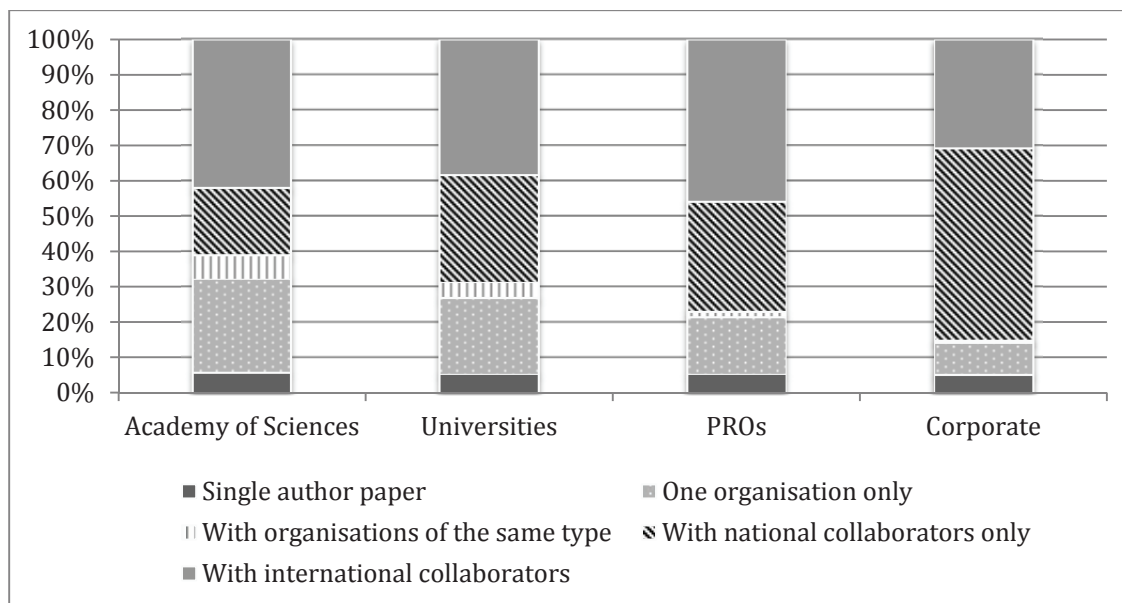


Figure 3. Institutional Diffusion of Russian Research System.

About two-fifths of academic publications are written either by a single author, or by a group of authors within RAS, and only 19% are collaborated with other Russian organisations. An international orientation is evident for PROs: over 46% of publications are internationally collaborated, but only 1.5% of publications are collaborated with other PROs. University organisations stand in the middle and have larger share of nationally collaborated publications than the Academy or PROs.

Weaknesses in international orientation and a reluctance to engage in national collaborative research projects is a particular concern for the Russian Academy of Sciences given that it dominates much of the Russian research system. In some RAS institutes, domestic collaboration rates with others outside of the home institute are noticeably low, for example just 11.6% in the Institute of Theoretical Physics RAS n.a. the Landau Institute of Theoretical Physics.

Conclusion

This exploratory study highlights three major path-dependent structural features of the Russian research system that are evident in Russia's nanotechnology research and publication activities. These structural features tend to be under-emphasized in other quantitative and qualitative studies, including those undertaken from within Russia itself. The available studies tend to focus on underfunding, deteriorating equipment, brain drain and other factors that, without a doubt, are very important in understanding the position of Russian science. In this research note, using bibliometric analysis in the case of nanotechnology, we draw attention to other less explicit but nonetheless important underpinning factors that frustrate the successful implementation of science and innovation policies and which may weaken returns on research investment. Reflecting upon and revising institutional practices of research that have remain largely unchanged since the breakup of the Soviet Union is an important challenge for Russian science policy. Some reform efforts have begun, but much more is likely to be needed to support the next generation of researchers.

Acknowledgements

This work was supported by the Economic and Social Research Council [grant number ES/J012785/1] as part of the project *Emerging Technologies, Trajectories and Implications of Next Generation Innovation Systems Development in China and Russia*.

References

- Ananyan, M. (2005). Nanotechnology in Russia: from laboratory towards industry. *Nanotechnology Law & Business*, 2, 194.
- Appelbaum, R.P., Parker, R., Cao, C., (2011). Developmental state and innovation: nanotechnology in China. *Global Networks*, 11, 298-314. doi:10.1111/j.1471-0374.2011.00327.x
- Bhattacharya, S. & Bhati, M. (2011). China's emergence as a global nanotech player: Lessons for countries In *Transition. China Rep.* 47, 243-262. doi:10.1177/000944551104700401
- Clemens, E.S. & Cook, J.M. (1999). Politics and institutionalism: Explaining durability and change. *Annual Review of Sociology*, 25, 441-466. doi:10.1146/annurev.soc.25.1.441
- Gokhberg, L., Fursov, K., & Karasev, O. (2012). Nanotechnology development and regulatory framework: The case of Russia. *Technovation*, 32, 161-162. doi:10.1016/j.technovation.2012.01.002
- Graham, L.R. (1998). *What Have We Learned about Science and Technology from the Russian Experience?* Stanford University Press.
- Ioffe Physical Technical Institute. (n.d.) Fizika Tvyordogo Tela / Physics of the Solid State Retrieved June 3, 2015 from: <http://journals.ioffe.ru/ftt/>
- Karaulova, M., Shackleton, O., Gok, A., Kotsemir, M.N., & Shapira, P. (2014). Nanotechnology research and innovation in Russia: A bibliometric analysis (SSRN Scholarly Paper No. ID 2521012). Social Science Research Network, Rochester, NY.
- Klochikhin, E.A. (2012). Russia's innovation policy: Stubborn path-dependencies and new approaches. *Research Policy*, 41, 1620-1630. doi:10.1016/j.respol.2012.03.023
- Liu, X., Kaza, S., Zhang, P. & Chen, H. (2011). Determining inventor status and its effect on knowledge diffusion: A study on nanotechnology literature from China, Russia, and India. *Journal of the American Society for Information Science and Technology*, 62, 1166-1176. doi:10.1002/asi.21528
- Liu, X., Zhang, P., Li, X., Chen, H., Dang, Y., Larson, C., Roco, M.C., & Wang, X. (2009). Trends for nanotechnology development in China, Russia, and India. *Journal of Nanoparticle Research*, 11, 1845-1866. doi:10.1007/s11051-009-9698-7
- Lux Research. (2013). *Nanotechnology Update: Corporations Up Their Spending as Revenues for Nano-enabled Products Increase.*
- Novoselov, K.S., Geim, A.K., Morozov, S.V., Jiang, D., Katsnelson, M.I., Grigorieva, I.V., Dubonos, S.V., & Firsov, A.A. (2005). Two-dimensional gas of massless Dirac fermions in graphene. *Nature*, 438, 197-200. doi:10.1038/nature04233
- Schiermeier, Q. (2007). Russia pins its hopes on "nano". *Nature*, 448, 233-233. doi:10.1038/448233a
- Springer. (n.d.) Physics of the Solid State - Springer Retrieved June 3, 2015 from: <http://www.springer.com/materials/journal/11451>
- Terekhov, A.I. (2011). Providing personnel for priority research fields (the example of nanotechnologies). *Herald of the Russian Academy of Sciences*, 81, 19-24. doi:10.1134/S1019331611010047
- Terekhov, A.I. (2012). Evaluating the performance of Russia in the research in nanotechnology. *Journal of Nanoparticle Research*, 14, 1-17. doi:10.1007/s11051-012-1250-5
- Terekhov, A.I. (2013). Russia's policy and standing in nanotechnology. *Bulletin of Science, Technology & Society*, 33, 96-114. doi:10.1177/0270467614524127
- Zhou, P. & Leydesdorff, L. (2006). The emergence of China as a leading nation in science. *Research Policy*, 35, 83-104. doi:10.1016/j.respol.2005.08.006